[illegible]ket 27554
in library
SD[illegible]
1/2/74

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 79, No. 1 **January 1972**

Articles

Cognitive Basis of Language Learning in Infants **John Macnamara** 1

A Neural Timing Theory for Response Times and the Psychophysics of Intensity **R. Duncan Luce and David M. Green** 14

Cognitive Control **Kenneth R. Hammond and David A. Summers** 58

Timing and Discrimination of Shock Density in Avoidance . . . . **John Gibbon** 68

Theoretical Note

Negative as well as Positive Synaptic Changes May Store Memory. . . **Mark R. Rosenzweig, Kjeld Møllgaard, Marian C. Diamond, and Edward L. Bennett** 93

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

NORMAN H. ANDERSON
ALBERT BANDURA
WILLIAM F. BATTIG
ROBERT A. BJORK
LYLE E. BOURNE
DALBIR BINDRA
N. JOHN CASTELLAN

ROBYN DAWES
J. A. DEUTSCH
EBBE EBBESEN
LEON FESTINGER
LEWIS R. GOLDBERG
KENNETH R. HAMMOND
JAMES HINRICHS

WALTER KINTSCH
CONRAD G. MUELLER
JACOB NACHMIAS
RAYMOND NICKERSON
DONALD NORMAN
M. FRANK NORMAN
KARALYN PATTERSON

MILTON ROKEACH
HARRIS B. SAVIN
RICHARD M. SHIFFRIN
GEORGE SPERLING
LARRY SQUIRE
EWART THOMAS
LEDYARD R. TUCKER

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts submitted to the *Psychological Review*. In each issue the Editor will list in the *Review* those editorial consultants who have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowledged.)

The *Psychological Review* is devoted to articles of theoretical significance to any area of scientific endeavor in psychology. Except for occasional articles, manuscripts exceeding 7,500 words will be discouraged. Ordinarily, manuscripts which consist primarily of original reports of research or of reviews of the literature should be submitted to other journals. Methodological and statistical articles are not appropriate for the *Psychological Review*. See the editorials in the *Review* (1965, 72, 1–2, 1971, 78, 1–2) for further details.

Beginning with the 1971 issue, Letters to the Editor will be published on an experimental basis. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts.* Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obtained from the APA Central Office.

All manuscripts must be submitted in triplicate. In the reference list, give journal titles in full; do not abbreviate. Figures should be submitted as glossy prints; duplicate figures may be photographic or pencil-drawn copies. Authors should retain a copy of the manuscript in case there is loss in the mail.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Manuscripts should be addressed to the Editor, George Mandler, Department of Experimental Psychology, South Parks Road, Oxford OX1 3PS, England.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of materials copyrighted by the American Psychological Association: Written permission must be obtained for copying or reprinting any tables, figures, or text of more than 500 words in length. Persons or organizations who obtain permission to reproduce any such copyrighted materials must include on the first page of the reproduced materials the exact copyright notice which appeared on the Association's original publication. There will be a charge of $10.00 per page for such permission except when the requester is the author of the article to be reprinted, or when the article is to be reproduced in a limited number of copies solely for instructional purposes.

HAROLD P. VAN COTT
Managing Editor

LORRAINE BIROS
Technical
Editor

HENRY S. SALMON
Advertising and Business
Manager

ANNE REDMAN
Subscription Manager

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1200 Seventeenth St., N. W., Washington, D. C. 20036

$10.00 volume

$2.00 issue

Second-class postage paid at Lancaster, Pa.

Send all communications, including address changes, to 1200 Seventeenth St., N. W., Washington, D. C. 20036. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PSYCHOLOGICAL
REVIEW

Diary No. 727(6)
Date 42
... No. ...
...eau Edul Psy Research.

Vol. 79, No. 1 JANUARY 1972

37

COGNITIVE BASIS OF LANGUAGE LEARNING IN INFANTS

JOHN MACNAMARA[1]

McGill University

BUREAU OF EDUCATIONAL AND PSYCHOL... LIBRARY GOVT. OF W. BENGAL CALCUTTA

The main point of the paper is that infants learn their language by first determining, independent of language, the meaning which a speaker intends to convey to them, and by then working out the relationship between the meaning and the expression they heard. The assumptions on which the thesis rests are discussed, the main one being that a speaker's linguistic system and his intentions are distinguishable. Evidence in support of the thesis itself is adduced at three levels: lexicon, syntax, and phonology. An attempt is made to describe the strategies, nonlinguistic as well as linguistic, by which children use meaning to decipher language.

The thesis of this paper is that infants learn their language by first determining, independent of language, the meaning which a speaker intends to convey to them, and by then working out the relationship between the meaning and the language. To put it another way, the infant uses meaning as a clue to language, rather than language as a clue to meaning. This thesis is a strong form of the theory that meaning plays an important part in language learning. Its most respectable antecedent is Brown's (1958) "original word game." I would like, however, to make use of recent developments in linguistics and psycholinguistics in developing the theory, and I would also like to present a theory which, unlike earlier ones, integrates vocabulary, syntax, and phonology.

The theory relates mainly to the beginning of language learning. Obviously once he has made some progress in language, a child can use language to develop his thinking. I am not, therefore, casting doubt on the belief that our thinking is enriched or brought into focus by language. Further, I am prescinding from the more general question of the extent to which our thinking is conditioned by the linguistic environment in which we live, and from problems relating to advanced levels of language, in particular certain aspects of poetic expression.

The approach to language learning which I am proposing had for some time been neglected in linguistics and in psycholinguistics. Part of the reason for this was the uneasiness with which behaviorism handled semantics. Even more important, however, was the impact of structuralism in linguistics which, following a suggestion of Bloomfield's (1933), sought to settle all grammatical questions without appeal to meaning. The advent of transformational grammar, in many respects so critical of linguistic structuralism, did nothing to change this situation. Indeed, in *Syntactic Structures*, Chomsky (1957, Ch. 9) went a

---

[1] Many of the ideas in this paper come from discussion with Nancy Katz of the McGill psychology department whose doctoral work is in this area. I am particularly grateful to her. I also owe much to Dick Tucker, Rose-Marie Weber, Charles Taylor, Roy Wright, Harry Bracken, Michael Corballis, all of whom either were or are my colleagues at McGill, and to Gillian and David Sankoff of the University of Montreal.

Requests for reprints should be sent to John Macnamara, Department of Psychology, McGill University, Montreal, Canada.

Bureau of Edul & Psyl. Research
(S. C. E R T) T 839

little further than Bloomfield and seriously questioned the view that meaning could be an aid to the discovery of grammar. The paucity of references to meaning in the literature on child language during the 10 years that follow the appearance of *Syntactic Structures* is scarcely a coincidence (for overviews see McNeill, 1966a, 1966b). In the late 1960s, however, grammarians came to rely more on semantics than before. Chomsky (1968) uses semantics to counter several arguments against the grammar which he had proposed; the arguments against him came from a group of transformational grammarians, such as Lakoff (1968), Lakoff and Ross (1967), Fillmore (1968), and McCawley (1968), who largely abolished the distinction between grammar and syntax. This turn of events, whatever the ultimate fate of the contending grammars with which it is associated, has engendered a readiness to seek a basis for language learning in infants among nonlinguistic cognitive principles. Perhaps the earliest and so far the most important study of child language in the new key is Bloom (1970). (A general swing in this direction can be seen in several books and articles, e.g., Brown, 1970; Ervin-Tripp, 1970; Kernan, 1969; McNeill, 1970; Slobin, 1971.)

While most of what has appeared on child language has focused on the production of speech, I would like to focus on the comprehension of speech. Some recent papers have indeed studied comprehension (Clark, 1971; Chomsky, 1969; Donaldson & Balfour, 1968). These, however, deal with children well beyond the initial stages of language learning. I know of only a very few papers which deal with the ability of two-year-old infants to comprehend speech (Shipley, Smith, & Gleitman, 1969; Smith, 1970). I would like to get behind these fine studies and to draw attention to the cognitive strategies which an infant brings to the task of comprehending speech. Apart from some very general remarks on such strategies (see, e.g., Chomsky, 1965, pp. 31 ff.; Fodor, 1966), the topic has been quite neglected. It is, however, a topic which is basic to the whole examination of language learning and, moreover, one which, as I hope to show, permits us several possibilities for empirical research. But first, some remarks on meaning.

I will use *meaning* to refer to those intentions which a speaker wishes to express in language, and following the lead of Frege (1952), I will regard sentences as names for the intentions which they express. Meanings can be assertions, negations, commands, and questions; they may include a reference to the speaker's physical environment, to his feelings (of pain, pleasure, uneasiness, etc.), and to his ideas or concepts. They may also include his attitudes of belief or doubt toward his assertions. In other words, meaning refers to all that a person can express or designate by means of a linguistic code. The code on the other hand consists of a set of formatives and syntactic devices whose main function is to relate meaning to the phonological system of the language.

I prefer to leave the problem of meaning at that, simply because I do not believe that we are able to define it more precisely without losing it. Perhaps the clearest attempts of behaviorists to treat meanings as mediating responses are those of Osgood (1963) and Mowrer (1960). With Fodor (1965), however, I regard all such attempts as unfortunate and misguided. Equally unfortunate and misguided, in my opinion, are the attempts of positivists such as Carnap (1956) and Quine (1960, 1961) to eliminate as far as possible talk about meaning and to replace it with talk of things present to the senses. All such attempts seem to me to purchase clarity by impoverishing human intelligence and by ignoring the dynamism of the act of understanding. Further than this, there is no need to go for the purposes of the present paper.

In stressing the distinction of meaning and linguistic code, I do not wish to imply a functional separation of thought and language. This would end up in what Quine (1960, p. 206) calls the "fallacy of subtraction," that is, the fallacy that meaning is what is left over when language has been filtered out. The fallacy can be avoided if with Saussure (1959, pp. 111 ff.) we regard an utterance as the embodiment of thought in language. Saussure points out that without thought, language is useless, and he also claims that thought without language is nebulous. While not going so far as to say that without language all thought is nebulous, I find Saussure's general thesis reasonable. Meaning and the linguistic code are best treated as though they were elements of a compound, much in the way that oxygen and hydrogen are the separate elements which combine to form molecules of water. That is, the two are not usually experienced separately, though they are distinguishable.

The thesis that infants use meaning as a clue to the linguistic code rests on the assumption that in the period when infants begin to learn language, their thought is more developed than their language. Support for this assumption is to be found in the work of Vygotsky (1962), Piaget (1963), and Sinclair-de-Zwart (1969), who adduce evidence that the development of thought is at first independent of language. They emphasize that at about the age of one, when an infant begins to understand language, he has already made many observations on the world about him, on himself, and on his own activities: he has succeeded in classifying many things and has seen the relationship between, for example, his own activities and the movement of various objects.

The thesis also has the support of the findings on cognition and affect in the deaf (for an overview see Vernon, 1971). Though there are difficulties in the interpretation of this work, it seems reasonably clear that persons who are severely retarded in language development nevertheless reveal all the essentials of human thought processes and human affect. It follows that thought and affect are distinguishable from language and that they can develop without benefit of language. It also follows that initially they can develop, independent of language, and outpace the growth of language in normal children.

Evidence from Vocabulary

At first sight it might appear that to learn the vocabulary of his language is the simplest linguistic task which confronts the infant. It might appear that all he has to do is to relate objects in the external world with their names. However, this illusion of simplicity soon disappears when we reflect on the vocabulary itself. Not only are there numerous names and descriptive terms which can be applied to a single object, but it soon transpires that descriptive terms refer to objects only through the speaker's complicated network of concepts (see Odgen & Richards, 1938). But how is the infant to match words and concepts?

In *Word and Object*, Quine (1960) argues convincingly, it seems to me, that no empirical test or set of tests could ever convince a reasonable but sceptical linguist that what appeared to be equivalent terms in two languages were in fact equivalent in meaning. The example he gives is of a linguist in a strange country who hears what seems to be the word for a rabbit. How is he to know whether the word might not mean "fur" or "head" or "animal" or "rabbitlike shape" or indefinitely many other things? Since, however, linguists make such decisions all the time and since it would be obtuse to deny that they are highly successful, it follows that they do not behave like Quine's sceptical linguist. They must rely on cognitive strategies which they bring with them to the task of describing a language.

Quine's arguments can be applied (though my use of them would not have his support) with few changes to the infant in the process of learning vocabulary. Like the adult linguist, the child's success depends on a set of cognitive strategies which function as shortcuts in the task of relating symbols to a speaker's intentions. That there must be such a set is evident from the fact that neither in the linguistic code nor in the physical environment are there sufficient constraints to explain how an infant could have even the most superficial success. For example, children do not form bizarre concepts to include foot and floor and exclude all else. There are quite marked constraints on what will be grouped together in a concept. Unfortunately the study of these cognitive constraints, which at the perceptual level interested the Gestalt psychologists, has been largely neglected, so most of what can be said about them is of a very tentative nature.[2]

It is obvious that an infant has the capacity to distinguish from the rest of the physical environment an object which his mother draws to his attention and names. It seems clear too that in such circumstances he adopts the strategy of taking the word he hears as a name for the object as a whole rather than as a subset of its properties, or for its position, or weight, or worth, or anything else. The same strategy would serve equally well in the numerous cases where something happens which draws an infant's and other people's attention to a particular object and they use the occasion to mention the object's name. Indeed the literature records several errors which, on the reasonable assumption that adults are good at guessing what children mean, show that the child is using just such a strategy (e.g., Blount, 1969, p. 62). An example is to be found in Hoffman (1968) who recounts that at the age of 17 months, one infant whom she was studying used to refer to the kitchen stove as *hot*. The likely explanation is that her mother told the infant not to touch the stove saying "*it's hot*," and the infant took *hot* to be the object's name. The theory thus exemplified fits in with some facts. Children learn names for colors, shapes, and sizes only after they have learned names for many objects. That is, names for entities are learned before names for certain attributes.

However, names for at least some of the various states, conditions, and activities of entities come very early. *Daddy* can be *gone*, *sitting down*, or *clipping the hedge*. A *toy* can be *broken*. One can *open* a *book* or *close* it. It is not that a state like *sitting down* is more clearly delineated than a color like *red;* the referents for each category show enormous variability. But the categories differ in at least one important and highly relevant respect. States and activities such as those just mentioned are not permanent attributes of entities, whereas colors and shapes usually are. If there is a differential set in small children to attend to varying states and activities rather than unvarying attributes, we need look no further for an explanation for the order in which the corresponding terms are learned.

A further hypothesis is that the child will not learn the name for states or activities until he has firmly grasped the name for at least some entities which exemplify such states and activities. Thus the order of learning would be as follows: names for entities, names for their variable states and actions, and names for more permanent attributes such as color. Perhaps the best research strategy would be first to test this hypothesis on clear-cut examples of each class. If the results are favorable, an argument based on construct validity would subsequently enable us to determine whether a particular doubtful term referred to a variable state or to a permanent quality in the child's world. The age at which the child learns the term would be the clue.

Some of the earliest words an infant learns to understand are proper names which single out the named object's individuality, for example, the names of his brothers and sisters and the name of the family dog. Other names are shared by many objects and connote what these objects hold in common. There are syntactic consequences attached to the distinction; one cannot speak of *a Fido*, but one can speak of *a dog*. Yet it seems unlikely that the young infant can use such evidence to distinguish between the two types of names. On the other hand, it is not at all improbable that he should assume that certain objects such as spoons have only common names, while other objects such as dogs and people should have proper names.[3] He may at first treat common and proper names when referring to the same objects—*Fido is a dog*—as optional proper names. If so, he will learn to differentiate when he discovers that only the family dog is called *Fido*, whereas all dogs are called *dog*.

---

[2] In addition to the Gestalt work on perception, which is surely related to concept formation, there is the extensive work of the ethologists on innate release mechanisms in lower animals and the part played in such mechanisms by selective attention to certain aspects or events of the environment (see Miller, Galanter, & Pribram, 1960, Ch. 5). The impact of the ethological work is to increase the credibility of a theory which incorporates restraints on perceptual attention and on concept formation in humans.

Initial investigation suggests that infants treat higher order common names as collective nouns.[4] For example a two-year-old child will treat "toys" as a collective noun which refers to a group of objects. Thus, while allowing that a particular object is a *truck* or a *train*, he will deny that it is a *toy*. On the other hand, he knows well what is meant by his *toys*. Moreover his response—that the *truck* is not a *toy*—does not seem to depend on what the surrounding objects are. It is the same whether or not the truck is surrounded by other toys or by objects which are clearly not toys, like spoons and cups. Thus it seems as if, for the child, *toy* is not a hierarchical term which contains *trucks* as a subcategory. It remains to be seen how children get themselves out of this situation, but presumably it is by a process similar to that which enables them to distinguish between proper and common nouns.

The pronouns, too, might appear impossibly difficult if we did not presuppose on the child's part the ability to guess their semantic force. How else could he learn to say *I* when he is addressed as *you?* Indeed autistic children usually fail to make this reversal.

So far we have dealt with some of the aspects of learning words which have referents in the concrete world of the environment, but there are also words which seem to refer primarily to cognitive operators and only secondarily to physical objects. Some such words, for example, *and*, *or*, and *true*, are acquired very early. Slobin and Welsh (in press) record a 27-month-old child's correct and spontaneous use of *and*. Brown, Fraser, and Bellugi (1964) record *all* and *more* as well as *some* for a 24-month-old boy. Even if some of the instances when such words are used are of doubtful validity, others seem genuine enough. It is inconceivable that the hearing of a logical term should generate for the first time the appropriate logical operator in a child's mind. Indeed the only possibility of his learning such a word would seem to be if he experienced the need for it in his own thinking and looked for it in the linguistic usage about him.

Evidence from Syntax

One might imagine that an infant who addresses himself to the task of learning syntax would to begin with adopt a simple hypothesis. He might, for example, suppose that a single semantic relationship would always be expressed by means of a single syntactic device or structure. Conversely, he might adopt the hypothesis that any change in syntactic structure signaled a change in semantic structure. However, such strategies would probably be less of a help than a hindrance. Consider 1 and 2 in each of which the same message is expressed by two different surface structures:

1. a. Give the book to me.
   b. Give me the book.
2. a. My hair is black.
   b. I have black hair.

Consider, too, 3 in which there is even a change in the syntactic category to which *kiss* belongs without any change in meaning:

3. a. Give me a kiss.
   b. Kiss me.

On the other hand, 4 shows that a single syntactic device can signal several different semantic structures:

4. a. In a minute (time "when").
   b. In a minute (time "how long").
   c. In a box.
   d. In a temper.

---

[3] Indeed some preliminary observations made by Nancy Katz of small babies support the idea that hey do make this discrimination.

[4] The work referred to in this paragraph is being one by Nancy Katz.

The task of the child is to detect the structures of syntax and relate them to semantic structures. The preceding examples show that the two are related in a rather flexible manner. The consistencies and regularities of syntax are to some extent independent of those of semantics. How, then, does the infant manage to relate the two? In keeping with the main lines of this paper, my answer will be that he uses independently attained meaning to discover at least certain syntactic structures that are of basic importance.

While I imagine that this position has the intuitive support of many students of child language, it runs counter to the assumption once widespread among linguists that syntax can be, even must be, described without recourse to meaning. Though few linguists would subscribe to this assumption today, the matter is so basic to the present discussion that it merits more than a cursory dismissal. Consider 5:

5. a. The boy struck the girl.
   b. The girl struck the boy.

How is the infant to know that these two sentences are not stylistic variants like 1a and 1b, or like 2a and 2b? There is in fact no way to decide the matter unless one has independent access to the meaning, probably through observing what is happening at the time sentences such as 5a and 5b are uttered. This is to say that without recourse to meaning, one could not discover so basic a syntactic device as the use of word order in English to express the subject-predicate relationship. The decision to treat 1a and 1b as stylistic variants must depend on the child's appreciation that while books can be given to people, people cannot be given to books. Thus the child must use his nonlinguistic knowledge to arrive at the notions of direct and indirect object and at the alternative ways of signaling them in English.

Before going on to speculate about how a child might use meaning to unravel syntactic puzzles, there are two qualifications to be made. First, not all syntactic puzzles must be or can be solved by semantic means. For example, gender in Indo-European languages does not correlate with any nonlinguistic classification of objects. Further, the rule that determines the number of the verb in English on the basis of the number of the surface subject is surely a purely syntactic one. Thus, though 6a and 6b mean roughly the same thing, the number of the verb is different in the two sentences:

6. a. The boys strike the girl.
   b. The girl is struck by the boys.

It is quite apparent then, that the child must have great skill in detecting the syntactic regularities of his language whether or not they can be tied in some way to meaning. Something of the French child's skill in relating word endings to gender can be seen in Tucker (1967).

The second qualification relates to the difference between *discovering* the syntactic devices of a language and *stating* them. A syntactic device which could not have been discovered without the aid of meaning might nevertheless be stated without any reference to meaning. To illustrate, a person who was intent on distinguishing Fords from Chevrolets might begin by reading the tradenames which appear somewhere on every car. After a time, however, he might come to rely on the shapes of fenders, rear windows, etc. There would even be a certain advantage in switching from tradename to distinctive features, since the latter can be recognized at a greater distance. Similarly in the realm of language. For example, it is altogether improbable that a child could discover the difference between transitive and intransitive active verbs without the aid of meaning. The presence or absence of following noun phrases is not an adequate criterion, as the examples of 7 show:

7. a. John drove.
   b. John drove a car.
   c. John drove home.
   d. I walked.
   e. I walked home.

The child must appreciate at some cognitive level that the action of driving has an effect on "car" but not on "home," and that the action of walking does not have an effect on "home."

The grammarian, on the other hand, might distinguish the two sorts of verbs by whether or not they have passives. Actually, this test is inadequate (see Lyons, 1968, Ch. 8), but it will serve as an illustration. Let us now suppose that a list of verbs which have no passive forms has somehow been formed. It follows that one can tell whether or not a verb is intransitive by whether or not it appears on the list. There is no need to refer to meaning. The fact, then, that children need meaning to establish the two classes of verbs and to decide which class a particular verb belongs to does not preclude the possibility or usefulness of formal tests which make no reference to meaning.

But to return to the main line of the discussion, how does the infant use meaning as a clue to certain syntactic devices? The most likely avenue to explore here is the prior learning of vocabulary, principally nouns, verbs, and adjectives. It seems natural to suggest that children initially take the main lexical items in the sentences they hear, determine referents for these items, and then use their knowledge of the referents to decide what the semantic structures intended by the speaker must be. This is in fact the strategy implicit in my treatment of examples 1 to 7. Once the children have determined the semantic structures, their final task is to note the syntactic devices, such as word order, prepositions, number affixes, etc., which correlate with the semantic structures. Such a strategy will yield most of the main syntactic devices in the language.

It is important to note, however, that the establishment of different semantic relationships among the formatives of a sentence does not immediately establish different syntactic relationships. For example, the relationship between *boy* and *ball* in 8a is quite different from the relationship between them in 8b:

8. a. The boy kicked the ball.
   b. The boy liked the ball.

Yet in each, *ball* can be taken as the grammatical object of the verb. On the other hand, the relationship between *I* and *home* in 9a is different both syntactically and semantically from the relationship between them in 9b:

9. a. I like home.
   b. I walk home.

The fact that *home* is not the direct object of *walk* in 9b is indicated by the fact that most nouns which follow *walk* are preceded by *to* or *toward:* one *walks to town, to the table, toward the door*. Perhaps the child notes this fact and combines it with the semantic information to arrive at the conclusion that the syntax of sentences such as 9b is different from that of sentences like 9a. This is of course highly speculative, but two general conclusions seem warranted: (*a*) the child cannot discover many syntactic structures without the aid of meaning, and (*b*) the child is prodigiously skilled at noting the regularities in the language which he hears.

The hypothesis that prior knowledge of vocabulary items and their referents is basic to the learning of syntax suggests that we return for a moment to some of the ideas discussed in the section on vocabulary. There it was noted that children at one stage take a word like *toys* as a collective noun rather than as a higher order noun with subcategories, *truck* and *doll*. It follows that *toys* should always be plural for the child at that stage, and that while he will understand 10a he will either not understand 10b or treat it as equivalent to 10a:

10. a. Those are your toys.
    b. That is your toy.
    c. A truck is a toy.

It follows further that he will not understand sentences like 10c in which the same referent is placed in two classes at once. Some of these matters are testable, but to go further at the present time would be to court disaster.

So far in this section, we have been concerned with declarative sentences, but there are of course many other forms of speech act such as questions and commands.

How the child learns to categorize speech acts correctly is almost a complete mystery, except that his doing so also presupposes an ability to determine the nature of the act independent of its syntactic form. In this connection, two observations will have to suffice. Lieberman (1967, pp. 44 ff.) cites evidence for the existence of an innate tendency in babies to react positively to friendly tones of voice and negatively to angry ones. Perhaps repeated commands tend to assume an angry tone, and so an angry tone might come to be taken by the child as an indication that he is either to do something or stop doing it. Later in the same book (p. 132), Lieberman examines 14 languages and finds that in 12 of them, questions can be signaled by means of rising pitch. While rising pitch is not then a universal marker of questions, it seems to have very wide currency as such. Could it be that children tend naturally to take rising pitch as a marker that a question is being asked?

But apart from such suprasegmentals, it seems clear that there must be a set of universal signs, of face, physical gesture, and bodily movement which the child interprets correctly and thus among other things comes to distinguish among speech acts. The evidence for such a set is growing (see Vernon, 1971), but as yet little is understood about how they relate to the learning of language. This is surely an important area of research. In the meantime we will have to content ourselves with the observation that the different forms of speech act probably correspond to deeply rooted mental attitudes of asserting, seeking for information, wishing to have others perform certain acts, etc. Indeed, such attitudes must be either innate or develop almost without benefit of learning. Endowed with such a set of attitudes, the child will be on the lookout for means of expressing them, nonverbal as well as verbal, and in this way come to discriminate the forms of different speech acts.

These suggestions of how an infant might use independently attained meaning to discover syntax are of course tentative. Yet they have the general support of several of the findings about the *production* of speech. For example, the suggestion that prior knowledge of vocabulary is vital to the learning of syntax fits well with the initial stages of speech production. First come single words that, as Bloom (1970), McNeill (1970), and several others have pointed out, infants employ to express a variety of semantic structures. Thus the single word milk can express: "This is milk," "Is this milk?," "I want milk," "I want more milk," etc. This indicates a heavy reliance on vocabulary items without supporting syntax. Moreover, when the child comes to combine words into two-word sentences, he continues to omit many of the syntactic features which adults include, such as plural and possessive markers, tense markers, auxiliaries, articles, etc. Yet Bloom (1970) has argued very forcibly that he does so for lack of syntactic processing space rather than for lack of semantic sophistication, and she was the support of many scholars in the area (see, e.g., Brown, 1970; McNeill, 1970; Schlesinger, 1971).

The point is equally clear when we consider children's attempts to repeat adult sentences. Brown and Bellugi (1964) and Brown and Fraser (1964), for example, pointed out that when required to repeat sentences, small children fasten onto meaning, drop the function words such as prepositions, and convey the substance of the message by means of the principal words. Slobin (1971) examines the data from several studies of child language in many countries and many different languages, and claims that such "telegraphese" is one of the universals of language learning.

In this connection it is also interesting to note the several suggestions in the literature that parents pay little attention to their small children's grammar, and only correct them on points of fact and discipline (see Bellugi-Klima, cited in Cazden, 1969). This observation, which is repeated in Gleason (1967), in Brown, Cazden, and Bellugi (1969), and in Brown and Hanlon (1970), implies that children learn grammar incidentally while parents attempt to put their thoughts and behavior right. However, too much must not be made of this point since children can learn from their parents remarks only if they can understand them, which probably implies after the very earliest stages, some knowledge of adult syntax. Indeed Shipley et al. (1969) found that small children made good use of adult syntactic devices in understanding adult sentences even if the children did not employ such devices in their own speech.

There is one further support in the literature for the thesis that in language learning the thrust to learn syntax comes mainly from meaning. It is the observation that children acquire first those grammatical devices which have a clear semantic force, such as numerous uses of tense and number markers, and only later acquired those which, like gender, have not (see Slobin, 1966). Indeed Slobin (1971) erects this into a tentative universal:

> Rules related to semantically defined classes take precedence over rules relating to formally defined classes.

I only know of one argument against the general thesis of this section. In an otherwise stimulating paper, Bever (1970, pp. 305–306) argued against the primacy of semantics in language learning. He presented sentences such as 11a and 11b to small children of varying ages:

11. a. The mother pats the dog.
    b. The dog pats the mother.

He required them to act out the meanings of the sentences with toys and found that relative to their performance on sentences like 11a, two-year-olds performed better than older children on sentences like 11b. That is, the two-year-olds seemed to be less upset by the improbability of the second type of sentence. Bever argues:

> The implication of this is to invalidate any theory of early language development that assumes that the young child depends on contextual knowledge of the world to tell him what sentences mean independent of their structure [p. 306].

This is a perplexing interpretation, seeing that in fact the younger children had interpreted the improbable sentences correctly more often than one might have expected. After all they were doing just what they were told. Besides, in absolute terms, the older children gave more correct responses to the improbable sentences. Indeed all sorts of explanations for the results other than the one given by the author come readily to mind.

Evidence from Phonology

The relationship between speech sounds and meaning is quite as loose as that between syntactic structures and meaning. A single speech sound either on its own or in combination with others frequently signals many different meanings, while just as frequently a variety of speech sounds signal a single meaning. Examples of the first sort of complexity are well known and beloved by punsters. There is *meat* and *meet*, or *watch* in the sense of "look at" and *watch* in the sense of "time piece." The other type of complexity is less well known and is worth spending some time on.

Phonological theory distinguishes between two major sorts of variation in the speech sounds of a language, one which is accompanied by variation in meaning and one which is not. Languages vary considerably, however, in their rules for assigning a particular variation to one category rather than another, and so the infant has to *learn* which is which. Examples of variation in sounds which are not accompanied by variation in meaning are many of the regional variations in the pronunciation of English or even the peculiarities of individual speakers. An example of a variation in sound which in English conveys a difference in meaning is that between *ship* and *sheep*. The interest of this example is that Spanish speakers have great difficulty in noticing the change because in Spanish it is of no importance.

I will give one further example to underscore the complexities in phonology which an infant may encounter and the need for access to meaning in order to master them. In modern Irish, the initial consonant of the word *bád* (boat) is modified after the definite article as indicated in 12:

12. Case	Normal script	First consonant
Nom.	*An bád*	[b]
Gen.	*An bháid*	[v]
Dat.	*An mbád*	[m]

What is noteworthy is that elsewhere in Irish, initial [m] and [b], for example, are contrasted as can be seen from the pair, *moladh* (praise) and *boladh* (smell). However, after the article in the dative case the initial [b] of *boladh* is replaced by [m] so that it sounds identical with *moladh*. We have, then a variation in sound which is sometimes accompanied by a variation in meaning and sometimes not.

In all, we have discussed three types of phonological variation, one accompanied by change of meaning, one not accompanied by change of meaning, and one which is sometimes accompanied by change in meaning and sometimes not. We began this section with changes in meaning without any change in sound. How is the infant to find his way through this many branched puzzle?

The answer must be that the infant is able to relate sound and meaning because he is able to tell what the speaker is speaking about independent of the speaker's language. This follows naturally from the evidence presented. The infant cannot use the sound system of a language to get at the meaning until he has learned the system. There seems to be little in the sound system itself which could direct the child in the choice of which sound differences to attend to and which to ignore. However, if he can determine the meaning and has sufficiently acute powers of auditory discrimination to begin with,[5] the infant can use meaning as the key to the sound system and thus arrive at the most elementary building blocks of his language.

Roger Brown (1958, pp. 213 ff.) describes a highly relevant experiment in his chapter on the original word game. He showed chips of different colors to native English speakers; each color was "named" with a nonsense syllable. Certain syllables differed only in vowel length which is not phonemic in English. The subjects were asked to group the chips in accordance with their names. Brown found that length of vowel was ignored if the corresponding chips differed little in color, but if the difference in color was marked, then different groups of chips were formed to correspond to the long and short vowels. In other words, subjects used the physical properties of the objects named as clues to the phonological units employed in naming them. Though his subjects were adults, the conclusion can be extended to infants without undue hesitation.

The claim that the infant uses meaning as the key to phonology needs to be filled out with details of how he does so, but unfortunately we know very little of the process. My remarks here are intended mainly as suggestions for further investigation. Let us begin by returning to the Irish example. Clearly the infant must use the fact that the referent is constant to arrive at the appreciation that [bád], [váid], and [mád] are all different realizations of the same phonological entity. In other words there are three different surface realizations of the same underlying morphophonemic form. Similarly, he must use the fact that *moladh* and *boladh* refer to different things (at least in equivalent environments) to avoid the mistake of treating these two words as different realizations of the same morphological entity. But while such a strategy works well for the examples I gave, how is the infant to avoid applying the constancy-of-referent strategy indiscriminately with disastrous phonological consequences? For example, *Fido* and *dog* might well be used to refer to a single animal, but to treat them as different surface realizations of the same phonological deep structure would be fatal. *Little* and *small* have more or less the same meaning, but they are phonologically quite distinct. The solution is that the various realizations of the Irish word *bád* are instances of a very general and simple phonological regularity. The rule for the dative (somewhat simplified), for example, is to voice the initial consonant if it is unvoiced and if it is voiced to replace it with the homorganic nasal (i.e., the nasal most closely related from an articulatory point of view). Part of the child's capacity to learn a language is his ability to detect such phonological regularities. Meaning alone cannot reveal them, though it probably directs the child to look for them. There are no phonological regularities relating *Fido* and *dog* or *little* and *small*. In the absence of such regularities the infant does not attempt to relate them phonologically despite the correspondences in meaning.

---

[5] There is now some evidence that infants are born with very fine powers of discrimination for speech sounds, and that learning for them consists mainly in ignoring differences in sound. The relevant material is well reviewed in Kaplan and Kaplan (in press).

To learn a language involves relating a system of intentions to a phonological system. To know the one is not ipso facto to know the other. Indeed the two are often related only through a complex system of syntax. I have had to have recourse to the syntactic system more than once in the foregoing discussion. For example, the phonological modifications of initial consonants in Irish are a function of syntactic rules which presumably must be learned together with the phonological modifications themselves. The combination of phonology and syntax are also presumably involved in the explanation of how the infant manages to distinguish such homonyms as *meat* and *meet*, and such nonhomonyms as *moladh* and *boladh*. This is hardly surprising in view of the fact that phonology and syntax are theoretical distinctions which are employed in the description of what is experientially a unit, the speech act. To explain the learning of phonology by means of syntax is not, however, to abandon the main thesis of this paper. In the previous section I argued that the learning of most basic syntactic relationships is possible only with the aid of meaning.

CONCLUDING REMARKS

I have treated of the learning of vocabulary, syntax, and phonology in separate sections, but the three could not be kept apart. It was necessary to base the learning of syntax in large part on knowledge of vocabulary, and it was necessary to base the learning of phonology in some part on knowledge of syntax. At this point it might be as well to correct the impression that infants use meaning to learn vocabulary, vocabulary to learn syntax, and syntax to learn phonology. It was not my intention to propose anything so neat. It is evident that some syntactic learning cannot be based on a knowledge of the referents of the lexical items employed. Similarly, phonology cannot be entirely dependent on prior knowledge of syntax. To argue that it was would be to argue in a circle, because without some independent grasp of phonology one could not perceive either words or syntactic devices. One cannot at once make the perception of vocabulary and syntax dependent on knowledge of phonology and make knowledge of phonology dependent on knowledge of vocabulary and syntax. The way out of this circle is to recognize that the interdependencies are only partial, though it is important to appreciate that at present they are quite obscure.

I have continually insisted on the child's possessing nonlinguistic cognitive processes before he learns their linguistic signal. By this I do not intend to endow the infant at birth with a complete ready-made set of cognitive structures. I accept Piaget's thesis that children gradually develop many of the cognitive structures, which they employ in association with language. Neither do I suggest that the child has a complete set of cognitive structures at the moment when he begins to learn language. All that is needed for my position is that the development of those basic cognitive structures to which I referred should precede the development of the corresponding linguistic structures. Since the acquisition of linguistic structures is spread over a long period, there is no reason that the acquisition of the corresponding nonlinguistic ones should not also extend well into the period of language learning.

I have repeatedly compared the infant to a linguist engaged in describing a language which to him is new. This seems a fair comparison provided the limp is recognized. Clearly, the infant does not produce explicit rules whereas the linguist does. The linguist aims to make explicit the native speaker's latent knowledge. Otherwise the tasks of infant and linguist are similar. Hockett (1961) points out that in fact no linguist has ever succeeded in deciphering a language without making use of meaning. He instances the Rosetta Stone which carried the same message in three languages, Greek, Demotic, and Hieroglyphic Egyptian. Because they knew Greek, Young and Champollion succeeded in deciphering the other two languages which no one at the time was able to read or understand. Similarly, Old Irish was deciphered by a Bavarian scholar named Johanne Kasper Zeuss. He worked on seven Latin manuscripts which carried glosses in Old Irish explaining the text. His great work, *Grammatica Celtica*, appeared in Latin in 1853, the first grammar of Old Irish.

It is not too fanciful to think of the infant as treating the sentences he hears as glosses on the events that occur about him. The grammar he writes is not in Latin or in any other language, but in some neurological code of which as yet not a single letter has been deciphered.

REFERENCES

BEVER, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and language learning*. New York: Wiley, 1970.

BLOOM, L. *Language development: Form and function in emerging grammars*. Boston: M. I. T. Press, 1970.

BLOOMFIELD, L. *Language*. New York: Holt, Rinehart & Winston, 1933.

BLOUNT, B. G. *Acquisition of language by Luo children*. Berkeley: Language Behavior Research Laboratory, University of California, 1969.

BROWN, R. *Words and things*. New York: Free Press, 1958.

BROWN, R. *Stage 1. Semantic and grammatical relations*. Cambridge, Mass.: Harvard University, Department of Psychology, 1970. (Mimeo)

BROWN, R., & BELLUGI, U. Three processes in the child's acquisition of syntax. *Harvard Educational Review*, 1964, **34**, 133–151.

BROWN, R., CAZDEN, C. B., & BELLUGI, U. The child's grammar from I to III. In J. P. Hill (Ed.), *1967 Minnesota symposium on child psychology*. Minneapolis: University of Minnesota Press, 1969.

BROWN, R., & FRASER, C. The acquisition of syntax. In U. Bellugi & R. Brown (Eds.), The acquisition of language. *Monographs of the Society for Research in Child Development*, 1964, **29**, 43–79.

BROWN, R., & HANLON, C. Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley, 1970.

CARNAP, R. *Meaning and necessity*. (2nd ed.) Chicago: University of Chicago Press, 1956.

CAZDEN, C. B. The psychology of language. In L. E. Travis (Ed.), *Handbook of speech, hearing and language disorders*. (Rev. ed.) New York: Appleton-Century-Crofts, 1969.

CHOMSKY, C. *The acquisition of syntax in children 5 to 10*. (Res. Mono. No. 57) Cambridge, Mass.: M. I. T. Press, 1969.

CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.

CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge, Mass.: M. I. T. Press, 1965.

CHOMSKY, N. *Deep structure, surface structure, and semantic interpretation*. Bloomington: Indiana University, Linguistics Circle, 1968. (Mimeo)

CLARK, E. V. On the acquisition of the meaning of *before* and *after*. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 266–275.

CROMER, R. F. "Children are nice to understand": Surface structure, clues for the recovery of a deep structure. *British Journal of Psychology*, 1970, **61**, 397–408.

DONALDSON, H., & BALFOUR, G. Less is more: A study of language comprehension in children. *British Journal of Psychology*, 1968, **59**, 461–471.

ERVIN-TRIPP, S. Structure and process in language acquisition. In, *Report of the twenty-first annual round table meeting on linguistics and language studies*. Washington, D. C.: Georgetown University Press, 1970.

FILLMORE, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart & Winston, 1968.

FODOR, J. A. Could meaning be an r_m? *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 73–81.

FODOR, J. A. How to learn to talk: Some simple ways. In F. Smith & G. A. Miller (Eds.), *The genesis of language*. Cambridge, Mass.: M. I. T. Press, 1966.

FREGE, G. *Philosophical writings*. (Ed. by Peter Geech & Max Black) Oxford: Basil Blackwell, 1952.

GLEASON, J. B. Do children imitate? Paper presented at the international conference on oral education of the deaf, Lexington School for the Deaf. New York, June 1967.

HOCKETT, C. F. Linguistic elements and their relations. *Language*, 1961, **37**, 29–53.

HOFFMAN, M. Child language. Unpublished master's thesis, McGill University, Department of Psychology, 1968.

KAPLAN, E. L., & KAPLAN, G. A. Is there any such thing as a pre-linguistic child? In J. E. Lot (Ed.), *Human development and cognitive processes.* New York: Holt, Rinehart & Winston, in press.

KERNAN, K. T. The acquisition of language by Samoan children. Unpublished doctoral dissertation, University of California, Berkeley, Department of Anthropology, 1969.

LAKOFF, G. Instrumental adverbs and the concept of deep structure. *Foundations of Language,* 1968, 4, 4–29.

LAKOFF, G., & ROSS, J. R. *Is deep structure necessary?* Cambridge, Mass.: Author, 1967.

LIEBERMAN, P. *Intonation, perception and language.* Boston: M. I. T. Press, 1967.

LYONS, J. *Introduction to theoretical linguistics.* Cambridge: Cambridge University Press, 1968.

MCCAWLEY, J. D. The role of semantics in a grammar. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory.* New York: Holt, Rinehart & Winston, 1968.

MCNEILL, D. Developmental psycholinguistics. In F. Smith & G. A. Miller (Eds.), *The genesis of language.* Cambridge, Mass.: M. I. T. Press, 1966. (a)

MCNEILL, D. The creation of language by children. In J. Lyons & R. J. Wales (Eds.), *Psycholinguistics papers: The proceedings of the 1966 Edinburgh Conference.* Edinburgh: Edinburgh University Press, 1966. (b)

MCNEILL, D. *The acquisition of language: The study of developmental psycholinguistics.* New York: Harper & Row, 1970.

MILLER, G. A., GALANTER, E., & PRIBRAM, K. H. *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston, 1960.

MOWRER, O. H. *Learning theory and the symbolic process.* New York: Wiley, 1960.

OGDEN, C. K., & RICHARDS, I. A. *The meaning of meaning.* New York: Harcourt, Brace & World, 1938.

OSGOOD, C. E. *Psycholinguistics.* In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 6. New York: McGraw-Hill, 1963.

PIAGET, J. Le langage et les opérations intellectuelles. In, *Problèmes de psycho-linguistique.* (Symposium de l'association de psychologie scientifique de langue française.) Paris: Presses Universitaires de France, 1963, 51–72.

QUINE, W. V. *Word and object.* New York: Wiley, 1960.

QUINE, W. V. *From a logical point of view.* Cambridge, Mass.: Harvard University Press, 1961.

SAUSSURE, F. DE. *Course in general linguistics.* New York: McGraw-Hill, 1959.

SCHLESINGER, I. M. Production of utterances and language acquisition. In D. Slobin (Ed.), *The ontogenesis of grammar.* New York: Academic Press, 1971.

SHIPLEY, E. F., SMITH, C. S., & GLEITMAN, L. R. A study of the acquisition of language: Free responses to commands. *Language,* 1969, 45, 322–342.

SINCLAIR-DE-ZWART, H. Developmental psycholinguistics. In D. Elkind & J. H. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget.* New York: Oxford University Press, 1969.

SLOBIN, D. I. Acquisition of Russian as a native language. In F. Smith & G. A. Miller (Eds.), *The genesis of language.* Cambridge, Mass.: M. I. T. Press, 1966.

SLOBIN, D. I. Universals of grammatical development in children. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics.* Amsterdam: North Holland, 1971.

SLOBIN, D. I., & WELSH, C. A. Elicited imitation as a research tool in developmental psycholinguistics. In C. A. Ferguson & D. I. Slobin (Eds.), *Readings on child language acquisition.* New York: Holt, Rinehart & Winston, in press.

SMITH, C. An experimental approach to children's linguistic competence. In J. R. Hayes (Ed.), *Cognition and the development of language.* New York: Wiley, 1970.

TUCKER, G. R. French speaker's skill with grammatical gender. Unpublished doctoral dissertation, McGill University, Department of Psychology, 1967.

VERNON, M. Language development's relationship to cognition, affectivity and intelligence. Paper presented at the meeting of the Canadian Psychological Institute, St. John's Newfoundland, June 1971.

VYGOTSKY, L. S. *Thought and language.* Cambridge, Mass.: M. I. T. Press, 1962.

(Received August 27, 1970; revision
received July 12, 1971)

Psychological Review
1972, Vol. 79, No. 1, 14–57

A NEURAL TIMING THEORY FOR RESPONSE TIMES AND THE PSYCHOPHYSICS OF INTENSITY[1]

R. DUNCAN LUCE AND DAVID M. GREEN[2]
Institute for Advanced Study, Princeton, New Jersey
University of California, San Diego

This psychophysical theory involves the following fundamental assumptions. At a hypothetical neural decision center, signal intensity is represented by several independent, parallel Poisson processes, whose rates are the same power function of physical intensity. All decisions about signal intensity are based on the observed times between successive neural pulses. The total number of these times observed per channel is at the option of the subject, up to the size of a memory store which is emptied when a decision is made. Overall response time is the sum of the decision latency, which depends both on the signal intensity and the decision rule, and a residual latency which is only assumed to be bounded. Decision rules are suggested for discrimination, recognition, magnitude estimation, detection, and simple reaction time designs, and predictions are derived from the theory in these cases and compared with existing data. Various familiar generalizations, such as Weber's and Bloch's laws and the inverse relation between reaction time and intensity, derive naturally from the theory. Crude estimates of all parameters—the exponent of the power function, the number of parallel channels, the size of the buffer store, and the bound on the residual times—are provided for sound intensity; estimates from different experimental designs appear to be reasonably consistent.

Elsewhere we have begun to evolve and test a detection and response time theory which, at least for pure tones of different intensities, seems to have merit (Green & Luce, 1967, 1971, 1972; Luce, 1966; Luce & Green, 1970). In this paper we formulate the theory more generally and show how it applies to a variety of psychophysical and reaction time (RT) designs.

The theory is based on three fundamental postulates which we now state; they are discussed more fully below.

1. The output of a sensory system, such as the ear, is treated as a set of sequences of discrete, brief events—"neural" *pulses*—on parallel channels with the following properties: (*a*) The duration of each pulse is negligible; (*b*) the times between successive pulses—the *interarrival times* (IATs)—on each channel and between channels are independent random variables (RVs); (*c*) when the intensity of a signal is constant, the IATs have a common distribution with an expected value that is a strictly monotonic decreasing function of the signal intensity.

We make two subsidiary assumptions which are conceptually quite secondary but are mathematically convenient and, apparently, are approximately correct. First, when the signal intensity is constant, these stochastic processes are assumed all to be *Poisson processes* with the same intensity parameter μ.[3] This means that on any one channel, the common distribution of the independent IATs is exponential and so, by 1*c*, its intensity parameter μ is a strictly increasing function of the signal intensity. Second, we shall assume that the latter function is, in fact, a power function. As we shall see, at least three different empirical results support this assumption.

2. Decisions as to when and/or what to respond are determined by decision rules that can be stated in terms of certain IATs.

---

[1] This research was supported in part by a grant from the Alfred P. Sloan Foundation to the Institute for Advanced Study and in part by the National Institutes of Health, Public Health Service, United States Department of Health, Education, and Welfare, and in part by the National Science Foundation. We wish to acknowledge the helpful comments of A. A. J. Marley and E. Thomas.

[2] Requests for reprints should be sent to David M. Green, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

[3] Undoubtedly, the values of the parameters differ in different channels. Our idealization is obtained by replacing the actual parameter values by average ones.

The principle underlying these rules (derived from 1*c*) is that any IAT is an estimate of (a monotonic function of) the signal intensity. We deal in this paper with only the simplest rules in which either a fixed number of IATs are observed or observations are continued until the first IAT occurs that is less than a criterion. Such models we call *timing models*. The time until a decision is reached is called the *decision latency*.[4]

One feature of the decision center is that it can store only a limited number of IATs at any one time. Sometimes, the observed behavior is limited by this storage capacity; other times it is limited by signal duration, which in effect limits the (expected) number of IATs.

3. The *residual latency*, which is defined as the overall RT less the decision latency, is an RV that is bounded and is independent of the decision latency.

Relation to Counting Models

The analysis of human sensory systems in terms of Poisson processes has become increasingly popular (Barlow, 1956, 1957; Bouman, 1961 [see references given there]; Creelman, 1961; Grice, 1968; Kohfeld, 1969; McGill, 1963, 1967; McGill & Goldberg, 1968; Siebert, 1965, 1968, 1970). The appeal is twofold: First, the observed all-or-none character of neural pulses invites description in terms of a continuous stochastic process. Second, of the possible processes, the Poisson is mathematically the most tractable and, as we argue below, is a tolerable approximation to some peripheral neural data. All of these other theories share the assumption that decisions are based not on IATs, but on a count of the Poisson events occurring within a prescribed time. The difference in our approach is exemplified by the problem of estimating an instantaneous heart rate. Either one can count the number of beats during a period of time and then calculate the average rate, or one can use the reciprocal of the time between the two successive beats spanning the instant of interest. The former estimate, which is usually based on a larger sample of IATs, is less variable, but because the average is constructed over some time period, it is not terribly sensitive to brief fluctuations in the heart rate. Since the detection of faint and/or brief signals that may arrive at any time is similar to noticing brief changes in the heart rate, we find the IAT decision rules appealing. In any event, timing models are worth investigating. Moreover, by emphasizing IATs rather than neural counters or accumulators, we avoid many of the elaborate and often ad hoc assumptions which are made concerning the memory for or integration of sensory information. It also has the distinct advantage of allowing one to derive the distribution of response times from the same assumptions as are used to arrive at the various stimulus–response probabilities, thereby permitting a more unified treatment of psychophysical and response time processes. Finally, certain empirical generalizations, such as Bloch's and Weber's laws, derive naturally.

Discussion of the General Theory

We repeat and discuss more fully our postulates. First, the sense organ, in our experimental studies the ear, is assumed to be a transducer that converts the intensity, $I(s)$, of the incoming fixed-frequency signal, s, into several trains of pulses with independent IATs. The intensity is represented by the temporal pattern of the pulses, not by their amplitude or spatial location; as a result, there is an inherent temporal ambiguity in knowing when changes of the input signal occur. In the case of constant intensity signals, we assume that the output on each of several channels is a Poisson process with intensity parameter $\mu(s)$ that, by assumption and confirmed by experiment, increases with $I(s)$. (This can be generalized to nonconstant intensities by assuming that the hazard function of the stochastic process is a strictly increasing function of $I(s)$.)

Since the pulses we are talking about are those on which decisions are based and hence are, presumably, located somewhere central in the brain, and since the properties most pertinent to our assumptions have only been carefully measured in neurons in the periphery, it is not clear whether the existing physiological data are relevant. According to Grossman and Viernstein (1961), Rodieck, Kiang, and Gerstein (1962), and Kiang (1965, especially pp. 95–101), successive IATs in the auditory nerves of cats are usually independent of each other, but their distributions are at best not quite exponential, and sometimes are decidedly not exponential. In the best cases, the observed distribution of IATs do not rise vertically at 0, but reach their peak value only after a few milliseconds, following which they are often well approximated by an exponential decay. The failure to find extremely short IATs is interpreted as refractoriness in the neurons. However, Rose, Brugge, Anderson, and Hind (1967; also see Brugge, Anderson, Hind, & Rose, 1969; Hind, Anderson, Brugge, & Rose, 1967; Rose, Brugge, Anderson, & Hind, 1969), using very refined techniques on the squirrel monkey, have shown that the distribution appears to be very nearly geometrically spaced at the period of the input signal. Thus, with a 1000-hertz (Hz.) tone, the IATs are approximately multiples of 1 millisecond (msec.) and have a distribution of the form $p^n q$. Similar frequency locked IAT data, with pulsed clicks and cats as subjects, were reported by Kiang (1965, Ch. 4). Of course, the exponential is the continuous approximation to the geometric and, provided that we do not care to discriminate as finely as 1 msec., it makes little difference which we use. Thus, although the Poisson assumption may not be quite correct, it seems to be a satisfactory first approximation in many cases. Ultimately, the exponential distribution can be replaced by, for example, a gamma distribution, but many fewer closed-form expressions will result. Two useful surveys of statistical analyses of single-unit activity are Kiang (1968) and Perkel and Bullock (1968).

---

[4] We follow the convention of using the word "time" for any time that we can observe directly and "latency" for any time that is an unobservable construct of a theory.

As we said, these data may not be relevant to our postulate. It should be noted that if a number of independent continuous stochastic processes with independent IATs are added, as may happen in parts of the nervous system, the resulting process tends toward the Poisson independent of the distribution of the original IATs.

There are relatively few data bearing directly on the dependence of $\mu(s)$ on $I(s)$, in particular, on the constancy of μ with constant I and on the discontinuous change in μ when we change the intensity discontinuously from one constant value to another, as in most psychophysical experiments. Kiang's (1968) data suggest that μ decreases over a period of minutes with a constant intensity signal, but over seconds the effect is negligible. He has also observed fibers in which $\mu(s)$ increases for a while with $I(s)$ and then diminishes as $I(s)$ is further increased.

As we shall see, psychophysical data indirectly suggest that for constant intensity signals,

$$\mu(s) = \begin{cases} \alpha[I(s) - I_0]^\gamma, & \text{if } I(s) \geq I_0 \\ 0, & \text{if } I(s) < I_0, \end{cases}$$

where α and γ are constants. Psychophysical data permit us to estimate γ, but not α. To anticipate, the exponent γ is obtained by magnitude estimation methods. S. S. Stevens (1970) marshals evidence to support the view that neural firing rates and other physiological measures of neural activity are power functions of intensity.

The second major component of the theory is a decision center that processes the sequences of pulses in some fashion to decide when to respond and/or which response to make. Since the IATs are independent and, with a constant intensity signal, identically distributed (actually, exponentially distributed under the Poisson assumption), it is plausible to assume that all decisions can be formulated in terms of the values of certain IATs. Exactly how decisions are reached depends both on inherent features of the nervous system and on the exact experimental situation, including our instructions and information feedback. By imposing different extreme experimental conditions, various limitations of the nervous system will become dominant, and we thus hope to discover what they are. Our tactic will be to generate a variety of models[5] as we consider the observable consequences of various decision rules. The common feature of all of the rules described in this paper, however, is that decisions are based on comparisons of means of IATs with each other or with a criterion; thus, the mean IAT plays a role somewhat comparable to the likelihood ratio RV in the theory of signal detectability (TSD; see Green & Swets, 1966).

The basic principle underlying all of the decision rules is that each IAT is an estimate of the signal intensity; this stems from the assumption that the expected value of the IAT is a monotonically decreasing function of the signal intensity (under the Poisson assumption, an IAT is an estimate of $1/\mu(s)$). With this principle in mind, three simple decision rules arise depending on the experimental design.

First, there are experiments in which the onset of a (possible) signal is well defined, and the problem is to identify which signal is being presented (included are yes–no detection, recognition, and magnitude estimation). In such designs, one or more IATs can be observed under the prior knowledge that they all arise from the same (Poisson) process. The rules that we shall examine have, at most, three free parameters: the number, J, of parallel channels (nerve fibers) over which intensity information is transmitted; the number, κ, of IATs per channel that a subject collects before initiating a response, and (sometimes) a criterion, β, against which the mean of the IATs is compared. When the subject is under considerable time pressure, for example, in certain RT experiments and, as we argue below, whenever the signals are both faint and brief, we shall assume $\kappa = 1$. When there is no time pressure, the value of κ must be estimated. It is clear that were the subject free to make κ arbitrarily large, he could, at the expense of taking enough time, make arbitrarily fine discriminations, which empirically is impossible. This leads us to postulate the existence of a buffer store in which only a *limited number* of IATs can be held prior to a decision. The size of the buffer, K, can be estimated from psychophysical experiments using brief but intense signals, such as recognition and magnitude estimation experiments.

Complications arise because it is not always clear which decision rule to assume; even when we restrict ourselves to rules only involving IATs and we specify J and κ, several alternatives seem equally plausible. A particularly thorny problem is how to treat simple detection experiments when the signal is presented only for a brief time. Just how should pulses that occur outside the observation interval be handled? It is evident that the observer ignores those events that occur far outside the listening interval, but it is unclear how sharply the observer's timing information can be used to demark the listening interval. Furthermore, the effect of this uncertainty must vary with signal intensity.

In those experiments in which the signal is presented for some fixed duration, there is a real possibility that the subject simply samples some fixed interval in time and counts the pulses in that interval. This is the usual analysis given by the counting models. If this is the case, then latency information should reflect variation in the end of this interval and in other fixed delays, but should be independent of which signal was present in the interval. We work out two timing models. The first simply waits for a neural event to terminate the trial, even if that event occurs well outside the listening interval. The second model assumes that IAT information initiates a response if it occurs before a deadline, which is set at the end of the listening interval, and initiates a response at that time if not before. It may well be that details of the instructions and the payoffs of the experimental task determine whether the counting or timing model is correct in these experiments with fixed signal durations. In many fixed signal duration experiments, with no time pressure, the counting models probably provide the more plausible theory.

---

[5] We follow the convention of saying that a specific decision rule for an experiment coupled with the other postulates of the theory is a model for that experiment. The experimental rejection of a model will be interpreted as the rejection of the decision rule, not the theory. Of course, if a sufficiently large number of plausible rules fail, the theory itself becomes suspect.

Second, there are experiments in which two or more signals are presented at well-defined times during each trial (included are forced-choice and pair-comparison designs). A mean IAT can be observed for each signal interval, with some number $k = J\kappa$ of observations entering into each mean, and then the means compared in order, for example, to decide which interval contained the most intense signal.

Third, there are experiments in which neither the occurrence nor identity of the signal is at issue, but only its time of occurrence (included are RT designs with random foreperiods and the method of free response). Here the task of the subject is to notice, by observing IATs, when the signal intensity is changed,[6] and the natural rule is for the subject to establish a criterion β with which he compares a running average of κ successive IATs on J channels. When that running average becomes less than a criterion β, a detection response is initiated. By increasing κ, the reliability of the detection can be improved, but at the expense of delaying it long after the actual onset of the signal. As we shall see, such rules are difficult to study, and we work out (incompletely) only the case of $J = 1$, $\kappa = 1$ for the RT experiment with random foreperiods. A special argument is used to generalize the results to $J > 1$. Considerably more research is needed on this class of rules.

Of course, rules other than these three types are possible. One of the most obvious for some of the above designs are truncated sequential rules. We have not undertaken their study here primarily because of the great difficulty one has in obtaining explicit expressions for response time densities. Nevertheless, they should be explored in the context of this theory. Two papers relevant to optimal detection for somewhat different stochastic processes are Bather (1967) and Shiryaev (1963).

For experiments in which the signal onset is clearly signaled in another modality, another type of rule is possible which takes advantage of a property peculiar to the Poisson process. This is the fact that the time from an arbitrary random event, such as the signal onset, to the next pulse of the process is an RV with exactly the same distribution as that of the IATs. Thus, decisions become more efficient if these times are combined with the IATs. The whole paper can be reconstructed using this type of rule. We elected not to use it for two reasons. First, if the process is not strictly Poisson, then the time to the first pulse after an event does not have the same distribution as that of the IATs and, in general, it would be a mistake to treat it in the same way. Second, life has evolved in a world in which the potential onset of a signal is rarely clearly marked in any way—most signals go unnoticed unless they attract attention to themselves—and so we would be surprised if a mechanism would have evolved just to take advantage of the extra information available in experiments having trials.

It should be noted that in all cases where we use a criterion β, we assume that it is maintained at a constant value throughout an experimental run. This is almost certainly wrong as the existence of trial-by-trial sequential effects argues. Our strategy is to assume that the variability in β is sufficiently small so that we can bury it in the variability of the IATs without too gross an error. Later, if certain decision rules seem viable, we will have to use sequential data to separate IAT and β variability.

The present decision rules are wholly inadequate to handle frequency discrimination. One possible approach is to suppose that different frequencies activate different, overlapping sets of parallel channels, and the subject attempts to identify from the observed IATs which subset is activated.

---

[6] By the dictionary definition, "detection" is the ideal term for this class of experiments, but that word has long been appropriated for such special recognition designs as yes–no and two-alternative forced-choice in which one signal is of zero intensity.

For low frequencies, another approach is suggested by the data of Rose et al. (1967), namely, that frequency can be inferred from the discrete pattern of the IATs, which are multiples of the period of the input signal, but subject to some random variation.

Since the pulses are strung out in time, no matter which rule we postulate some time is consumed in arriving at a decision. In those situations with a well-defined signal onset, we can observe a response time, part of which is contributed by the latency of the sensory decision process. Because the expected value of the IAT increases with decreasing signal intensity, the decision latency and so the observed response time is magnified by using weak signals. In a sense, weak signals serve as our "microscope" to explore these elementary, but basic decision mechanisms of the brain.

We use the letter "l" to denote the generic density of sensory decision latencies, sometimes adding subscripts to identify which signal was presented and which response was made. Most of the mathematical work in the theory involves using the assumed decision rule and the other postulates of the theory to derive expressions for the various decision densities. These formulas include as parameters such things as the Poisson intensity parameters of the sensory output and, where appropriate, J, κ, and β.

The third major component of the theory is the sum of all times taken to complete the response excluding the time absorbed by the decision process. This includes, at least, the times (*a*) needed to transduce the energy into pulses, (*b*) required for a pulse to travel from the transducer to wherever decisions are made, (*c*) taken in travel from there to the locus of the response, and (*d*) needed for the muscles to act. This list may not be exhaustive. The sum of all these nondecision times we call the residual latency.[7] We assume that these residual latencies are continuous RVs, that they are independent of each other and of the decision latencies, and that they are bounded. We denote their (unknown) density by r, with a subscript referring to the response made if there are two or more. Clearly, changing either the mode of stimulation or the mode of response will change r. The boundedness assumption means that there is some $\tau > 0$ such that $r(t) = 0$ for all $t \geq \tau$.

We know and assume very little about r. We do not attempt to construct any theory of the residual times. From our point of view, they are a nuisance factor making it difficult to get at the decision process. Although postulating a bound on the residual latencies seems most plausible, it should be realized that it is sometimes crucial in our analysis. This bound, together with the exponential IATs, which are therefore unbounded, sometimes permits us to get a fairly clear look into aspects of the sensory decision process uncontaminated by the residual times. Were the residual times unbounded, we could not use the tails of observed distributions to estimate parameters.

If we denote by f_j the density of observed response times when response j is made to the signal, then by our independence assumption, it is the convolution of the decision and residual densities, l_j and r_j (see Figure 1), that is,

$$f_j(t) = \int_0^t l_j(x) r_j(t - x) dx. \qquad [1]$$

Because all of our densities are 0 for $t < 0$, the Fourier transform[8] of f_j is

$$F_j(\omega) = \int_0^\infty e^{-i\omega t} f_j(t) dt,$$

$$\omega \text{ is real and } i^2 = -1.$$

---

[7] In our previous publications we called it the "response latency" because, in thinking about free response designs, we had focused on the time from a decision to a response. In talking with others, we discovered that this was badly misleading, especially for designs with trials, and so it seems best to alter the terminology. The residual latency corresponds to what was called the "irreducible reaction time" in the classical RT literature.

[8] We follow the convention of using small Latin letters for densities and their capitals for the corresponding Fourier transform.

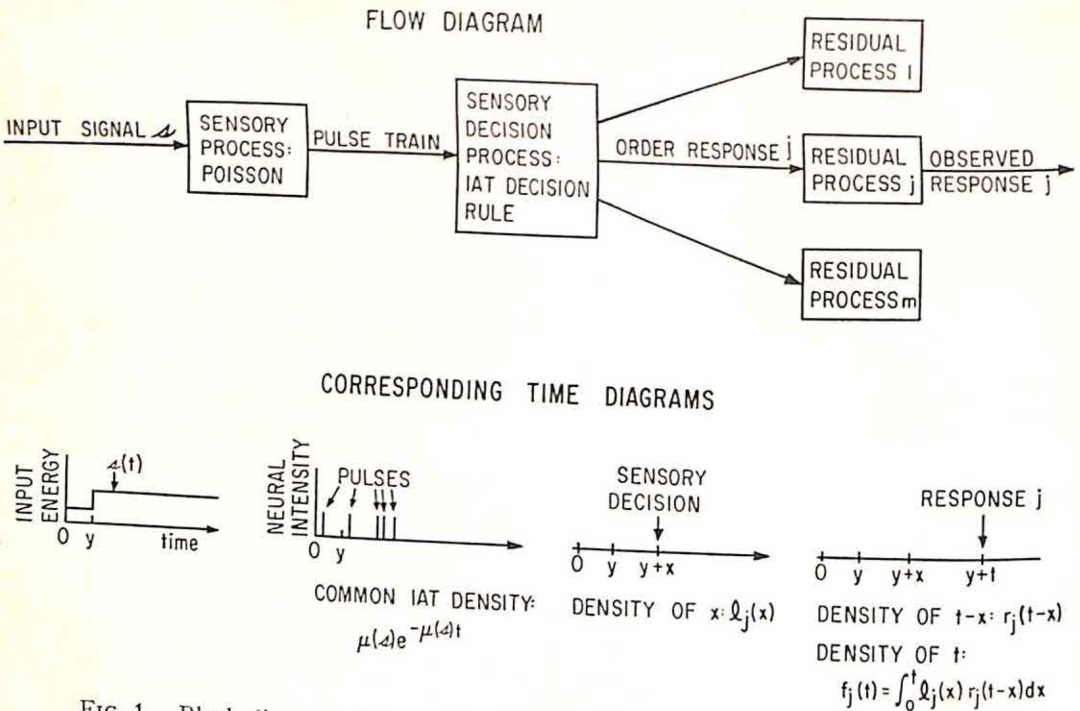


FIG. 1. Block diagram of model and time diagrams of the output at each stage.

It is well known that if f_j is a convolution of l_j and r_j, then[9]

$$F_j(\omega) = L_j(\omega)R_j(\omega). \qquad [2]$$

This suggests the following general procedure for testing a model. State the proposed sensory decision rule and, suppressing the response label j, derive L (usually, though not always, by deriving l and then computing its Fourier transform) as an explicit expression in terms of parameters, such as the Poisson parameters of the sensory process, J, κ, and β. Devise ways to estimate these parameters (we often use the boundedness of r to show that the shape of the tail of f depends only on l and not on r, which permits us to use the tails to obtain estimates). Calculate from the data $F(\omega)/L(\omega)$, which is an estimate of $R(\omega)$. By inverting this transform, we have an estimate of r. We can then see, first, whether it is a density and, if so, whether it is independent of manipulations of the parameters. If not, then doubt is cast on the decision rule. This was done successfully in Green and Luce (1971).

We show precisely how the boundedness of r can be used in parameter estimation. Suppose it has been shown that for sufficiently large t, l acts like a gamma distribution, that is,

$$l(t) \cong ct^m e^{-\eta t} \quad \text{for } t \geq \tau^*. \qquad [3]$$

Then by Equation 1,

$$\begin{aligned} f(t) &= \int_{t-\tau}^{t} l(x)r(t-x)dx \\ &\qquad [\text{since } r(t-x) = 0 \text{ for } t-x \geq \tau] \\ &\cong c\int_{t-\tau}^{t} x^m e^{-\eta x} r(t-x)dx, \\ &\qquad\qquad \text{for } t \geq \tau^* + \tau \\ &= c\int_0^{\tau} (t-y)^m e^{-\eta(t-y)} r(y)dy \\ &= e^{-\eta t}\sum_{i=0}^{m} a_i t^i, \end{aligned} \qquad [4]$$

where

$$a_i = c\binom{m}{i}(-1)^{m-i}\int_0^{\tau} y^{m-i}e^{\eta y}r(y)dy.$$

---

[9] Other classical transforms, including the Laplace, have the same property of converting convolutions into multiplications. We use the Fourier because there exists an excellent computer program, known as the Fast Fourier Transform, for numerically calculating transforms and their inverses.

For $m = 0$, f is proportional to $e^{-\eta t}$ if $t \geq \tau^* + \tau$, and so η can be estimated from the tail of f. For $m > 0$, the dominant term is $t^m e^{-\eta t}$ and again η can be estimated. For some models, but not all, l is shown to have the above form.

Another approach was suggested many years ago by Christie and Luce (1956). Suppose the experiment is run twice with signals of different intensities, but with the same response. Assuming that the residual latency is unaffected by the signal intensity, we have

$$\frac{F_1(\omega)}{F_2(\omega)} = \frac{L(\omega,\mu_1)}{L(\omega,\mu_2)}.$$

It is not clear how to estimate μ_1 and μ_2 or, if they are estimated, how to test goodness of fit. Note that in the special case where signal 2 is extremely intense, we can assume that the decision time is negligible, in which case F_2 estimates R and so

$$L(\omega,\mu_1) \to F_1(\omega)/F_2(\omega) \quad \text{as } \mu_2 \to \infty.$$

A major unsolved problem in this case is to think through the estimation of parameters, which is not quite conventional since the data have been heavily processed before we get to an estimate of L or l.

Organization of the Paper

The remainder of the paper simply applies the above general theory to a variety of psychophysical tasks: discrimination of an increment in intensity, recognition of one of several intensities, magnitude estimation, simple detection, and simple reaction time. We group these designs not by any operational criterion, but by the extreme theoretical distinction of whether the subject is free to use maximum κ permitted by his buffer store or is "forced" to use a $\kappa = 1$ decision rule. The former arises when there is no time pressure and the signals are both of sufficient duration and intensity, what we call "strong" signals, so that K IATs are very likely to occur on each channel during the presentation. In particular, those tasks, such as discrimination, recognition, and magnitude estimation, in which the signals are so intense that there is no problem whatsoever in detecting them, seem to permit the use of the maximum (see, however, the section entitled "Estimates of J and K" for a strong qualification of this remark). The latter is assumed to arise when there exists an experimentally imposed time pressure, which can take at least two forms: explicit instructions and payoffs to be as fast as possible or the use of weak signals of brief duration. We do not attempt here to study the difficult intermediate category of experiments in which no explicit time pressure exists, but the intensity-duration of the signal is insufficient, on the average, to fill the buffer store during the presentation of the signal.

We take up these topics in the order: strong signals, weak signals, and RT. The reason for this order is increasing analytic complexity. For strong signals, we derive only statements about choice probabilities and do not study the response times. For weak signals, we derive both probability and temporal relations, but the latter are comparatively simple. Finally, in the RT section we focus almost entirely on the complex problem of detecting the onset of a signal when it can occur at any time.

The basic procedure followed in the analysis of each design is the same. First, we state the exact nature of the experimental task. Next, we describe a decision rule which might be suitable. Finally, we derive equations for the densities of the various decision latencies and for the conditional response probabilities from the decision rule and other assumptions of the theory, calculate Fourier transforms of the densities, and indicate (if we know) how to estimate parameters. Wherever possible, we cite relations to well-known results, either theoretical or empirical.

Parameter Estimates

Problems of estimating parameters arise in any theory of this type, but they are especially acute in the present theory for several reasons. First, the parameters often enter the equations as products—$\mu(s)\beta$, JK, etc.—and so they are not completely identifiable. If $JK = 100$, then either $J = 100$ and $K = 1$ or $J = 1$ and

Bureau of Ednl & Psyl. Research

$K = 100$ is an equally good possibility. This ambiguity cannot be reduced if we stay in a single experimental context. Only by studying behavior in a variety of experiments can we arrive at estimates in which we have some confidence. Second, in many cases we do not have completely satisfactory analytic methods for estimating parameters. Partly this results from difficulties in solving equations; partly because we have not worked out the correct model. For example, we are reasonably confident from various data that $J > 1$; however, at present we have formulated only a $J = 1$ model for simple RTs with random foreperiods. Estimates from this model differ from $J > 1$ models for other experiments by as much as an order of magnitude. Finally, at present we are at the mercy of a few observations from different experiments run in a number of different laboratories for various reasons unrelated to our theory. Nevertheless, to help clarify the theory and to indicate how one might proceed, we make tentative estimates of the basic parameters from existing data. The reader should not take these estimates too seriously; we do not.

Comparison with Other Theories

As we have already suggested, the present theory has much in common with well-known theories; but it differs in significant ways. It is well to appreciate the similarities and differences.

Just as in TSD, we assume that the subject observes a sensory RV and compares it with a criterion or with another sensory RV. Here, however, the RV is one or more IATs, whereas in TSD it is a likelihood ratio, a rather sophisticated function of the sensory datum and the context of the experiment. Since our sensory process is assumed to be Poisson, the sum of $k = \kappa J$ IATs has a gamma distribution with k degrees of freedom, and so the estimate of μ, the reciprocal of the mean of k IATs, has a known distribution different from the normal generally postulated in TSD. Of course, the IATs are temporal RVs and permit us to derive latency distributions, whereas the likelihood ratios have nothing explicit to do with time. John (1967) has sketched a theory in which an RV, monotonic with time, is compared with a criterion.

Other theories, including TSD, the Poisson counting models, and some simple RT models in which there has been an explicit attempt to deal with time (Carterette, 1966; Edwards, 1965; Estes, 1960; LaBerge, 1962; Laming, 1968; Stone, 1960) differ from ours by assuming the existence of a temporal structure other than that created by the output of the sensory process. The main idea is that there is some sort of time quantum, either imposed by the experimental design (TSD) or by the subject, and within each quantum an RV is observed and compared with a criterion. In most such models, observations are taken until one, or the average, RV gets outside an interval of indecision, that is, a Wald-type sequential test is employed. Unfortunately, it is usually impossible to derive exact closed-form expressions for the resulting latency distributions. By contrast, we postulate no internal temporal structure other than that provided by the output of the sensory process, and the resulting IATs are our sensory RVs; often, latency distributions are easily derived. Of course, if there really is an inherent human time quantum over which information is integrated, as Kristofferson (1967a, 1967b) has argued, it is important to model it correctly; however, the current concensus is against that hypothesis and it seems imperative, therefore, to work out temporal models which do not include any time quantum. Ours do not include such a quantum.

PSYCHOPHYSICS OF STRONG SIGNALS

Assuming that a buffer store exists, we may give a theoretical definition of a strong signal as one of such intensity and duration that, almost certainly, more IATs will occur during its presentation than can be accommodated in the buffer store. More formally, if K is the (IAT) size of the store, T the duration of the signal, and μ its Poisson parameter, then the signal is strong if μT is considerably larger than K. For strong signals that are also intense (large μ), detection is not the issue; rather, experiments are aimed at understanding the apparent magnitude of the signals or the discriminability of two such signals when their intensities are about equal.

The models we shall investigate, called *limited memory timing*, assume that the subject is allowed, both by the signal and by the reward structure, to accumulate as many IATs as are permitted by his buffer store, namely $k = KJ$, where J is the number of distinct channels. On the average, this requires K/μ seconds per channel. We assume that he uses the average duration of the IATs as an estimate of $1/\mu$, and we derive some probability predictions of this model for discrimination, recognition, and magnitude estimation designs. We do not write down any of the equations for the densities of decision latencies because, with large μ, these latencies are such a small fraction of the response time that they probably cannot be used effectively to study decision rules. (If the present theory is conceptually correct, most simple RT studies, with their very intense signals, elicit information mainly about the residual latency, not the decision one.)

At least two other plausible models exist for strong signals. In the *counting models*, which have been investigated by others, no memory limitation is postulated, and the subject is assumed to count the number of pulses that occur within the signal interval. Just how that interval is demarked for the subject is a moot point. In addition, the counting model has the unfortunate feature of predicting error-free performance as $\mu T \rightarrow \infty$; it is well known that increasing T beyond, at most, a few seconds does not improve performance.

Perhaps the most appealing models from an intuitive point of view are those we may call *sequential, limited memory timing*. The subject is assumed to accumulate information until either the evidence in favor of one response becomes sufficiently persuasive or until the memory is saturated, whichever comes first. Their intuitive appeal is strongly offset by their mathematical complexity. It is often impossible to calculate the distribution of the number of IATs to a decision without making grossly simplifying assumptions.

Discrimination of Intensity

In this section we shall consider pair-comparison designs in which two signals, s_1 and s_2, differing only in intensity, are each presented for duration T on each trial. We let $\mu_i = \mu(s_i)$. The subject is to report which, the first or second, is the more intense in his judgment. The design in which one signal is held fixed and the other varied is called the method of constant stimuli.

We simply assume that as each signal is presented, the subject stores the first $k = KJ$ IATs, records the total time (i.e., forms their sum), compares the two totals, and reports whichever is smaller as the more intense signal. If we let $\mathbf{T}_i$ denote the total time RV arising from signal s_i, then

$$\begin{aligned} p(1 \mid s_1,s_2) &= P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0) \\ &= P(\mathbf{T}_2/\mathbf{T}_1 \geq 1) \end{aligned}$$

is the probability that the first signal is reported as more intense. It is well known (see Parzen, 1962, p. 137) that

$$p(1 \mid s_1,s_2) = \frac{(2k-1)!}{[(k-1)!]^2} \int_0^\rho [x(1-x)]^{k-1} dx, \quad [5]$$

where

$$\rho = \frac{\mu_1}{\mu_1 + \mu_2}.$$

This is the incomplete Beta distribution $I_\rho(k,k)$ or, equally, the F distribution

$$Q\left(\frac{1-\rho}{\rho}; 2k,2k\right) = Q\left(\frac{\mu_2}{\mu_1}; 2k,2k\right).$$

It can be written in another way,

$$p(1 \mid s_1,s_2) = \sum_{i=0}^{k-1} f(i; k,\rho)$$

where

$$f(i; k,\rho) = \rho^k \binom{k+i-1}{i} (1-\rho)^i$$

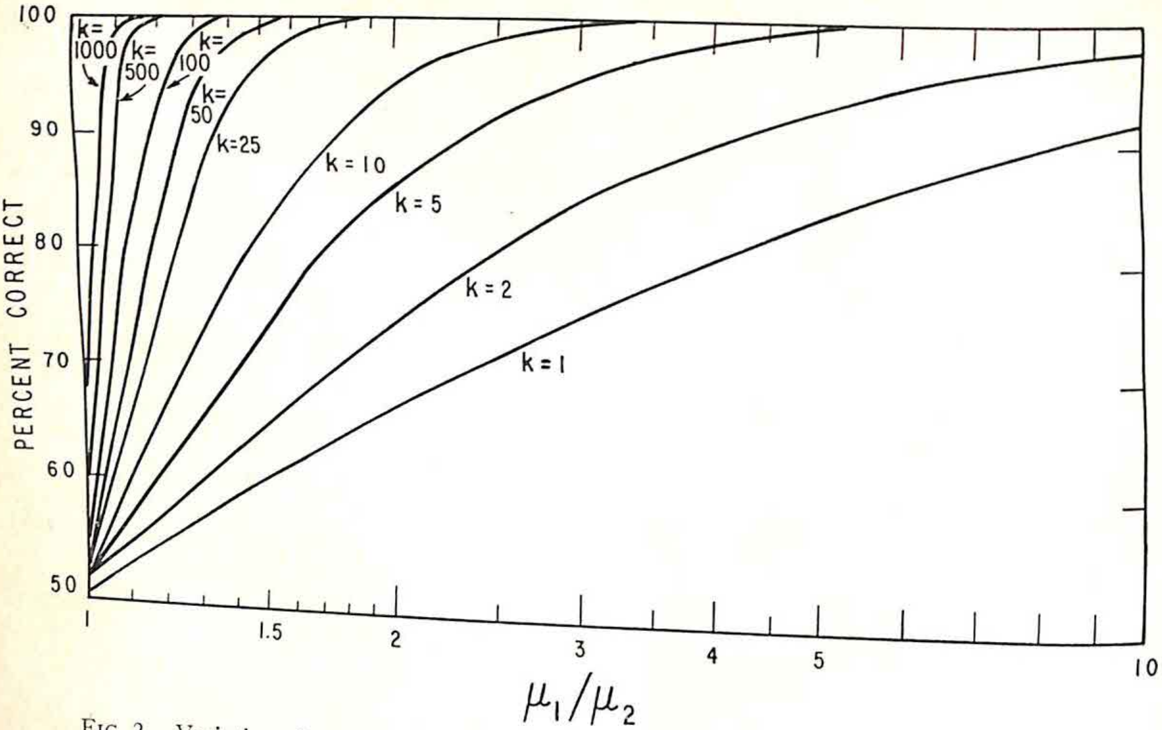


FIG. 2. Variation of percentage of correct response in an intensity discrimination task. (The abscissa is the ratio of the rate parameters for the two signals. The parameter is the number of independent channels on which the decision is based, $k = KJ$.)

is the negative binomial distribution. As Green and McGill (1970) note in a similar model, for large k, Equation 5 is well approximated by

$$p(1|s_1,s_2) \cong \int_{-\infty}^{d'/\sqrt{2}} \mathfrak{N}(0,1), \qquad [6]$$

where

$$d' = \sqrt{k - 1/2}\left(\frac{\mu_1}{\mu_2} - 1\right). \qquad [7]$$

Figure 2 shows $p(1|s_1,s_2)$ as a function of μ_1/μ_2 for several values of k.

Observe that $p(1|s_1,s_2)$ is a constant if and only if μ_1/μ_2 is constant. Since, by hypothesis,

$$\frac{\mu_1}{\mu_2} = \left[\frac{I(s_1) - I_0}{I(s_2) - I_0}\right]^\gamma = \left[\frac{I(s_1) - I(s_2)}{I(s_2) - I_0} + 1\right]^\gamma,$$

μ_1/μ_2 is a constant if and only if

$$\frac{I(s_1) - I(s_2)}{I(s_2) - I_0}$$

is a constant, so Weber's law holds for $I(s_2) \gg I_0$. As $I(s_2)$ approaches I_0, two effects contribute to the observed increase in

$$\frac{\Delta I}{I} = \frac{I(s_1) - I(s_2)}{I(s_2)}.$$

Most obviously, I_0 can no longer be neglected and $\Delta I/(I - I_0)$ is constant only if $\Delta I/I$ is increasing. More subtly, as μ is decreased, the number of IATs that can be observed before the signal is terminated must, necessarily, decrease. In effect, the size K of the storage is functionally decreased and, as we show in Appendix A, as long as

$$\frac{\mu_1}{\mu_2} > \frac{1 + 1/(2k-1)^{\frac{1}{2}}}{1 - 1/(2k-1)^{\frac{1}{2}}}$$

then in order to maintain $p(1|s_1,s_2)$ at a constant value, μ_1/μ_2 must be increased as k is decreased.

Another interesting feature of the model shows up if, instead of memory limitations, we assume extreme time pressure ($\kappa = 1$) (due to short signal duration) and a single channel ($J = 1$) and so set $k = 1$ in Equation 5,

$$p(1|s_1,s_2) = \rho = \frac{\mu_1}{\mu_1 + \mu_2}.$$

Notice that this generalizes nicely to m signals:

$$\begin{aligned} p(1|s_1,\cdots,s_m) &= \int_0^\infty \mu_1 e^{-\mu_1 t} \prod_{i=2}^{m} e^{-\mu_i t} dt \\ &= \int_0^\infty \mu_1 e^{-\sum_{i=1}^{m} \mu_i t} dt \\ &= \frac{\mu_1}{\sum_{i=1}^{m} \mu_i}. \end{aligned} \qquad [8]$$

Models based on Equation 8 have been called choice models and were investigated by Luce (1959) and others. Within the present framework, we would only expect them to hold for intensity discrimination when the subject is under severe time pressure and has only one channel.

Biased Discrimination of Intensity

If the payoffs are asymmetric, so that it matters to the subject which type of error he makes, then we anticipate biased responses. If his available information is the two random variables $\mathbf{T}_1$ and $\mathbf{T}_2$, the bias could take various forms. One possibility is to say signal 1 when $\mathbf{T}_1 - \mathbf{T}_2 \geq \epsilon$; another is to say signal 1 when $\mathbf{T}_2/\mathbf{T}_1 \geq \delta$. With $\epsilon = 0$ and $\delta = 1$, these are equivalent; in general, however, they are distinct. We consider them separately. For the first, the argument given in Appendix A yields

$$P(\mathbf{T}_2 - \mathbf{T}_1 \geq \epsilon) = \begin{cases} \rho^k \left[\sum_{i=0}^{k-1} \binom{k+i-1}{i} (1-\rho)^i \sum_{j=0}^{k-1-i} \frac{(\mu_2 \epsilon)^j}{j!} \right] e^{-\mu_2 \epsilon}, & \epsilon \geq 0 \\ (1-\rho)^k \sum_{i=0}^{k-1} \binom{k+i-1}{i} \rho^i \left[1 - e^{\mu_1 \epsilon} \sum_{j=0}^{k-1-i} \frac{(-\mu_1 \epsilon)^j}{j!} \right] \\ \qquad + P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0), & \epsilon < 0 \end{cases} \qquad [9]$$

where $P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0)$ is given by Equation 5. Observe that for $k = 1$, Equation 9 reduces to

$$p(1|s_1,s_2) = \begin{cases} \rho e^{-\mu_2 \epsilon}, & \epsilon \geq 0 \\ (1-\rho)(1 - e^{-\mu_1 \epsilon}) + \rho, & \epsilon < 0 \end{cases} \qquad [10]$$

from which the following receiving operator characteristic (ROC) curve derives:

$$p(1|s_1,s_2) = \begin{cases} \rho \left[\frac{p(1|s_2,s_1)}{1-\rho} \right]^{\mu_2/\mu_1}, & p(1|s_2,s_1) \leq 1 - \rho \\ 1 - (1-\rho) \left[\frac{1 - p(1|s_2,s_1)}{p} \right]^{\mu_2/\mu_1}, & p(1|s_2,s_1) > 1 - \rho \end{cases} \qquad [11]$$

Plots of the ROC curve generated by Equation 9 for several values of k are shown in Figure 3.

The other possibility yields (Appendix A)

$$P\left(\frac{\mathbf{T}_2}{\mathbf{T}_1} \geq \delta\right) = \frac{(2k-1)!}{[(k-1)!]^2} \times \int_0^{\rho(\delta)} [x(1-x)]^{k-1} dx \qquad [12]$$

where

$$\rho(\delta) = \frac{\mu_1}{\mu_1 + \delta \mu_2}.$$

For $k = 1$,

$$p(1|s_1,s_2) = \frac{\mu_1}{\mu_1 + \delta \mu_2}$$

from which we obtain the ROC curve

$$\frac{p(1|s_1,s_2)}{p(2|s_1,s_2)} = \left(\frac{\mu_1}{\mu_2}\right)^2 \frac{p(1|s_2,s_1)}{p(2|s_2,s_1)},$$

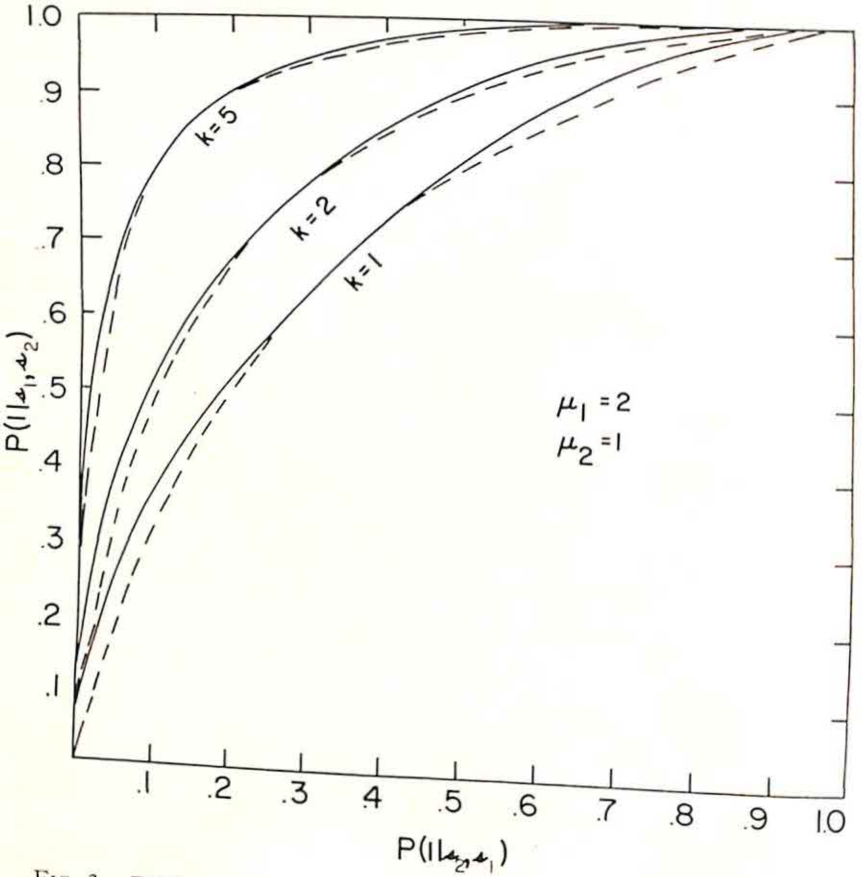


FIG. 3. ROC curves for intensity discrimination experiment with $\mu_1 = 2$, $\mu_2 = 1$ and two different decision rules. (The solid curves are generated by varying the criterion on the *difference* on the sum of k IATs associated with each interval. The dotted curves are generated by varying the criterion on the *ratio* of the sum of k IATs associated with each interval.)

which is identical to that of the choice model (see Luce, 1963a). The curves generated from Equation 12 for various values of k are also shown in Figure 3. It is unlikely that experimental data could be used to distinguish between the curves based on Equations 9 and 12.

Before leaving our discussion of the discrimination of intensity differences, we should comment on an obvious deficiency of models in this area. The present timing model says nothing about the interval separating the two signal presentations. According to the model, it is irrelevant, whereas, in reality, discriminability does diminish with increasing separation between the two signals. The obvious way to remedy this situation within the counting model is to assume that the count is not recalled with perfect fidelity (Kinchla & Smyzer, 1967). Similar assumptions might be made within the timing model concerning the fidelity of the time recorded for the first signal. Various alternatives are possible, but none has been worked through to any interesting predictions; we leave this as a conspicuous open problem.

Recognition of One of m *Signals*

By a recognition (absolute identification) design we mean that on each trial, one of m signals, $s_1, \cdots, s_m$, is presented and the subject is to identify which by using one of m responses that are in one-to-one correspondence with the signals. Usually the correspondence is established as being order preserving. In our analysis we assume the signals differ only in intensity, hence s_{i+1} is more intense than s_i for all i. Our approach to recognition parallels that of Thurstonian psychophysics, except we use a gamma distribution rather than the normal. In particular, we assume that the signal duration is sufficiently long and the intensities sufficiently great that, almost certainly, k IATs can be observed, and that their sum is a random variable $\mathbf{T}_i$ which has a gamma density with parameters k and μ_i when signal s_i is presented. We assume that the subject selects $m - 1$ criterion values

$$0 < \beta_{m-1} < \cdots < \beta_p < \beta_{p-1} < \cdots < \beta_1 < \infty,$$

and that his decision rule is to respond p if and only if

$$\beta_p \le \mathbf{T} < \beta_{p-1}.$$

Clearly, then

$$\begin{aligned} p(\mathbf{R} \le p|s_i) &= \int_{\beta_p}^{\infty} \frac{\mu_i{}^k t^{k-1} e^{-\mu_i t}}{(k-1)!}\,dt \\ &= \int_{\beta_p \mu_i}^{\infty} \frac{y^{k-1}e^{-y}}{(k-1)!}dy. \qquad [13] \end{aligned}$$

For large k the central limit theorem yields

$$P(\mathbf{R} \le p|s_i) \cong \int_{-\infty}^{z_i(p)} \mathfrak{N}(0,1), \qquad [14]$$

where

$$z_i(p) = \sqrt{k} - \frac{\mu_i \beta_p}{\sqrt{k}}. \qquad [15]$$

Thus, Equation 13 is approximately a nonequal variance Thurstonian model. The equal variance case, in which $z_i(p) = (\beta_p - \mu_i)/\sigma$, has been proposed for intensity recognition by Durlach and Braida (1969) and is tested experimentally in Braida and Durlach (1972).

Given recognition data, the $(m-1)m\ z_i(p)$ can be obtained from Equation 14 (the cases where $z_i(p) = -\infty$ or ∞ are omitted from the estimation scheme). We can then use Equation 15 in the following (nonoptimal) way to estimate parameters. We note that from Equation 15 it follows immediately that

$$z_i(p) - z_j(p) = (\mu_j - \mu_i)\beta_p/\sqrt{k}$$

and

$$z_i(p) - z_i(q) = \mu_i(\beta_p - \beta_q)/\sqrt{k}.$$

Thus, if the model is correct, the geometric mean of

$$\frac{z_i(p) - z_j(p)}{z_i(q) - z_j(q)} = \frac{\beta_p}{\beta_q}$$

over all $i,\ j,\ i > j$ yields an estimate of β_p/β_q. Similarly, the geometric mean of

$$\frac{z_i(p) - z_i(q)}{z_j(p) - z_j(q)} = \frac{\mu_i}{\mu_j}$$

over all $p,\ q,\ p > q$ yields an estimate of μ_i/μ_j.

Since $\mu_i > 0$ and the unit of measurement is not determined, there is no loss of generality in choosing the unit so that

$$\sum_{i=1}^{m} \mu_i = 1.$$

We then have as an estimate of μ_i,

$$\hat{\mu}_i = \frac{1}{m}\sum_{j=1}^{m}\left(\frac{\hat{\mu}_i}{\hat{\mu}_j}\right).$$

With that in hand, we can use Equation 15 to estimate $\beta_p/\sqrt{k}$ by forming the geometric mean of

$$\frac{z_j(p) - z_i(p)}{\hat{\mu}_i - \hat{\mu}_j}$$

over all $i,\ j,\ i > j$. Finally, from Equation 15 we see that the mean over all i and p of

$$\hat{\mu}_i\left(\frac{\beta_p}{\sqrt{k}}\right) + z_i(p)$$

is an estimate of $\sqrt{k}$.

With all of these parameters estimated, we may use Equation 15 to predict the $m(m-1)$ (less the ∞ cases) observed $z_i(p)$ from the $2m - 1$ free parameters. These predictions can be displayed as a scatter diagram and evaluated by a chi-square test.

It should be of interest to compare the estimate of μ_i/μ_j obtained in this way with those obtained by magnitude estimation (see the next section).

Magnitude Estimation

Experimentally, magnitude estimation (ME) and recognition have much in common: a single signal is presented on each trial, and a numerical response is required in which a larger number is assigned to a more intense signal. They differ, of course, in that the experimenter establishes a one-to-one correspondence in the recognition design between the responses "1," "2," $\cdots$, "m" and a fixed set of m signals; whereas he asks the subject in the ME design to reveal his "built-in" one-to-one correspondence between all positive real numbers and all possible signals for an unspecified sample of these signals. One notable consequence of this difference is that there are correct and incorrect responses in the recognition experiment, and so payoffs can be used, whereas no concept of error is appropriate in the ME design. Our theoretical analysis differs considerably.

We simply assume that the subject attempts to estimate $c\mu$, where c is an adjustable unit parameter, by observing k IATs and using the total time $\mathbf{T}$ as his information about the signal. In fact, we assume that he responds $\mathbf{M} = ck/\mathbf{T}$ as his estimate of the magnitude of the signal. Since $\mathbf{T}$ is distributed as a kth order gamma, its reciprocal has the Pearson Type V distribution (see Kendall & Stuart, 1958):

$$P(\mathbf{M} = x) = \frac{(kc\mu)^k e^{-kc\mu/x}}{(k-1)!x^{k+1}}. \qquad [16]$$

Perhaps the most unusual feature of this density is its high tail—of the form $x^{-(k+1)}$ as x becomes large. This agrees with the observations of Luce and Mo (1965) and Schneider and Lane (1963). In fact, the tails are so high that moments beyond $k - 1$ do not exist. The mode, mean (for $k > 1$), geometric mean (for $k > 1$), and variance (for $k > 2$), which are easily derived, are given by:

$$\left(c\frac{k}{k+1}\right)\mu, \quad \left(c\frac{k}{k-1}\right)\mu, \quad \left(cke^{C - \sum_{i=1}^{k+1} 1/i}\right)\mu,$$

$$\left(c\frac{k}{k-1}\right)^2 \frac{1}{(k-2)}\mu^2, \qquad [17]$$

respectively, where C is Euler's constant ($\cong .577216$). Since the observed mean or median magnitude estimates are, approximately, power functions of signal energy (S. S. Stevens, 1957), the postulate that μ is a power function of intensity is reasonable. Moreover, this gives a way to estimate the exponent γ. This assumption was used to account for Weber's law.

If this model is correct, it is clearly better not to work with the data and Equation 16 directly, but rather to form reciprocals. These have the density

$$P\left(\frac{1}{\mathbf{M}} = y\right) = \frac{(ck\mu)^k y^{k-1} e^{-ck\mu y}}{(k-1)!} \qquad [18]$$

which, because of its exponential tail, has much more stable moments than Equation 16. The mean and variance both exist and are

$$\frac{1}{c\mu} \quad \text{and} \quad \frac{1}{(c\mu)^2 k},$$

and so we have the following estimate of k:

$$\hat{k} = \text{mean}^2/\text{variance}. \qquad [19]$$

If all forms of signal intensity are recoded in pulse trains with a rate that increases with intensity, then comparisons of intensity between modalities is a possibility. At least intensities on different modalities are represented by a common measure, time, and so we need only postulate that the nervous system is capable of making comparisons of IATs originating in the different senses. Thus, the a priori mysterious, but successful, method of cross-modality matching makes complete sense in this theory. Because of our power function assumption, we predict a power function matching relation with an exponent that is the ratio of the two ME exponents, which is what is observed. In addition, since both signals exhibit variability in their IAT representation, we predict that the distribution of matches should be more variable than the reciprocal of the corresponding MEs. We do not know of data bearing on this prediction.

If we approach ME data much as we did recognition data, we have from Equation 16

$$\begin{aligned} P(\mathbf{M} \le p \mid s_i) &= \int_0^p \frac{(kc\mu_i)^k e^{-kc\mu_i/x}}{(k-1)!x^{k+1}} dx \\ &= \int_{kc\mu_i/p}^{\infty} y^{k-1} e^{-y} dy/(k-1)! \\ &\cong \int_{-\infty}^{z_i(p)} \mathfrak{N}(0,1), \end{aligned} \qquad [20]$$

where

$$z_i(p) = \sqrt{k}\left(1 - \frac{c\mu_i}{p}\right). \qquad [21]$$

Here p is a cut point established by us, not by the subject. Considering only the μ_i term, the dependence on k and the criterion p are inverse to that found in the recognition model (Equation 15).

Estimates of J *and* K

According to the ME model (especially Equation 17) and S. S. Stevens' (1957) data, we know that for 1000-Hz. tones, approximately,

$$\frac{\mu_1}{\mu_2} = \left(\frac{I_1}{I_2}\right)^{.3},$$

where I is acoustic power. It is also well known that the Weber fraction corresponding to $p(1|s_1,s_2) = .75$ is about .08, so the corresponding parameter ratio is

$$\frac{\mu_1}{\mu_2} = (1.08)^{.3} = 1.023.$$

Using Equation 6, the normal approximation to Equation 5, we have an estimate of $k = JK$ of about 1750.

With k known, it is sufficient to estimate either J or K. The only methods that have occurred to us are for J. One possibility, using ME data, is suggested by data of S. S. Stevens (1966) and J. C. Stevens and Hall (1966), which were both concerned with the dependence of MEs on signal duration T. For both brightness and loudness, the MEs increase linearly in log T and then abruptly become constant. In the case of brightness, the location of the corner is a function of intensity, being at about 10 msec. for an 85-decibel (db.) signal and at about 100 msec. for a 55-db. one. This strongly suggests that the subject does not fill his buffer store before determining his estimate, but rather that only a few IATs are collected on each channel. The times indicate a 10 to 1 ratio in the μs over a 30-db. range which agrees with the fact that the range is equivalent to a stimulus ratio of 10^3 and that the brightness exponent is about 1/3. Indeed, the numbers are such that on the basis of threshold data, using fixed intervals, we are probably safe in assuming that only one IAT is obtained from each channel.

The situation for loudness differs in one important respect: the corner, which is at 150 msec., is independent of intensity. This strongly suggests that the number of IATs accepted increases with intensity, which agrees qualitatively with the finding mentioned by Luce and Mo (1965) that their estimate of k increases; however, they reported a factor of only 4 over a 47-db. range whereas we would predict, using an exponent of .3, a factor of 26 in order for the location of the corner to remain fixed. In contrast Schneider and Lane (1963) using magnitude production, but converting their data to an equivalent estimation form, found the standard deviation of the estimates to be 1/5 of the mean. From Equation 17, this yields the estimate of $k = 27$.

There appears to be a discrepancy here in that the Stevens and Hall data for loudness estimation suggest that k increases with intensity, whereas those of Schneider and Lane for loudness production suggest it does not. Further investigation is required to determine if this is a difference between estimation and production or if something deeper is wrong with the theory.

Assuming only one IAT per channel, $\kappa = 1$, for the Schneider and Lane experiment, $J = 27$ and so $K = 1750/27 \cong 65$.

The reader is no doubt perplexed, as are we, about why 65 IATs should be collected per channel in a discrimination experiment and only 1 or a few in the ME procedure. The only explanation that we have thought of, and it is wholly ad hoc, is that calculating a reciprocal is considerably more complex than comparing two sums, and so vast amounts of storage capacity are reserved to carry it out. We would be more convinced of this had the loudness data been similar to the brightness data and so suggested $\kappa = 1$, independent of intensity. There seems to be a problem in need of better explanation.

PSYCHOPHYSICS OF WEAK SIGNALS

We turn next to one of the classical areas of psychophysics, the detection of weak signals. Here the focus continues to be on various response probabilities, but in contrast to strong signals the response times begin to assume some interest because the delays of the sensory-decision process are no longer negligible relative to delays of the residual process.

The essential problem for the subject is to decide on each trial whether or not a signal is present in a noise background. Because the signal is weak, we know that the density of Poisson events is low or, to put it another way, that the expected value of an IAT is moderately large. We may impose time pressure on the subject either by deadlines and payoffs or by instructions while leaving the signal on until a response occurs, as in a type of RT experiment with catch trials, or by using signals of such short duration that a failure to decide quickly almost certainly brings into play pulses not due to the signal, as in a conventional yes–no design. We shall suppose that both of these procedures invoke decision rules that involve only a single or very few IATs from each channel.

It is not customary to regard fixed-interval detection experiments as involving any strong time pressure, but clearly they do if anything like the present theory is correct. For example, suppose that a 500-msec. interval is specified and that a faint signal is presented with, let us say, a value of $\mu = 4$, then the expected time to the first event is 250 msec. and the second is 500 msec. This suggests that on the average, it would be futile to wait for a second IAT and that, indeed, it may be futile to wait for even one if the second impulse lies beyond the end of the signal interval.

We examine situations having only two signals: noise of constant power, n, and a signal of fixed frequency and intensity in that noise, s. Thus, there are only two Poisson parameters, $\mu = \mu(s)$ and $\nu = \mu(n)$, where $\mu > \nu$. We use obvious mnemonics for the responses: Y for yes, N for no, 1 for the first interval, 2 for the second. A typical conditional response probability is denoted $p(\mathrm{Y} | n)$, that is, the probability of a yes response given that no signal was presented. The response time and decision latency densities are subscripted similarly, so for example, $f_{n\mathrm{Y}}$ is the density of the yes response times to the presentation of no signal, and $l_{s\mathrm{N}}$ is the density of the no decision latencies to the presentation of a signal.

Yes–No Design with Response-Terminated Signals

It is convenient to begin with an unconventional yes–no design because it is one of the simplest to analyze. With that as an example, we turn later to the usual fixed-interval design. Consider an experiment in which the potential onset of a signal is clearly indicated to the subject (by an intense signal in another modality, e.g.). On some trials there is no signal. On the others, a constant intensity signal is presented and remains on until the subject responds Y or N. The subject is pressed (by payoffs or instructions) to respond as fast as he can. In other words, this is a simple RT experiment with zero foreperiod and catch trials, but with a faint response-terminated signal (which is not usual in the RT literature).

The decision rule we shall suppose is that the first κ IATs are obtained from each channel and if the average of these $k = \kappa J$ IATs is less than β, respond Y; otherwise respond N. The decision time is then simply the time of the slowest $(\kappa + 1)$ pulse, which has the $(\kappa + 1)$-order gamma distribution $\gamma_{\kappa+1,\mu}(t)$ with intensity parameter μ. Since if J RVs have a common density, that of the largest is

$$l_s(t) = \frac{d}{dt}\left[\int_0^t \gamma_{\kappa+1,\mu}(x)dx\right]^J$$

$$= J\gamma_{\kappa+1,\mu}(t)\left[\int_0^t \gamma_{\kappa+1,\mu}(x)dx\right]^{J-1}. \quad [22]$$

From this, a direct calculation yields the mean reaction time (MRT),

$$\mathrm{MRT}_s = \bar{r}_{\mathrm{Y}} p(\mathrm{Y}|s) + \bar{r}_{\mathrm{N}} P(\mathrm{N}|s) + \frac{h(J,\kappa)}{\mu}, \quad [23]$$

where

$$h(J,\kappa) = J\int_0^\infty \frac{x^{\kappa+1}e^{-x}}{\kappa!} \times \left[\int_0^x \frac{y^\kappa e^{-y}dy}{\kappa!}\right]^{J-1} dx \quad [24]$$

and $\bar{r}_Y$ and $\bar{r}_N$ are the mean residual latencies for, respectively, Y and N responses. For large κ, Equation 24 can be transformed by the central limit theorem into

$$h(J,\kappa) \cong \kappa + 1 + \sqrt{\kappa + 1}H(J),$$

where

$$H(J) = J\int_{-\infty}^\infty x\mathfrak{N}(0,1)\left[\int_{-\infty}^x \mathfrak{N}(0,1)\right]^{J-1} dx$$

is the mean of the largest of J normally distributed RVs with mean 0 and variance 1. Several values of $H(J)$ are given in Table 1.

To get some idea of the error of the approximation for small κ, we compute $h(2,1) = 2.750$ as compared with the approximate value of 2.797, an error of less than 2%.

Note that if $\bar{r}_Y = \bar{r}_N$, then if we can vary κ experimentally, Equation 23 implies

$$\mathrm{MRT}_n = \frac{\mu}{\nu}\,\mathrm{MRT}_s + \bar{r}\left(1 - \frac{\mu}{\nu}\right). \quad [25]$$

This linear regression provides a test of the model as well as a way to estimate μ/ν and $\bar{r}$.

Although the density of decision times is not simple in this model, the expression for the response probabilities is. Let **T** denote the sum of $k = J\kappa$ IATs. Because each IAT is exponentially distributed, **T** has a gamma distribution of order k. Thus,

$$\begin{aligned} p(\mathrm{Y}|s) &= P(\mathbf{T} \leq k\beta) \\ &= P\left(\frac{1}{\mathbf{T}} \geq \frac{1}{k\beta}\right) \\ &= \int_{\frac{1}{k\beta}}^\infty g_s(x)dx, \end{aligned} \quad [26]$$

where

$$g_s(x) = \frac{\mu^k e^{-\mu/x}}{(k-1)!x^{k+1}}, \quad [27]$$

which is a Pearson Type V density (see Kendall & Stuart, 1958).

TABLE 1

$H(J)$ VERSUS J

J	$H(J)$
2	.564
5	1.163
10	1.539
20	1.867
60	2.319
100	2.508
200	2.746
500	3.037
1000	3.241

Note.—These values are taken from Table I of Tippett (1925). Values of $H(J)$ for all values of J, $2 \leq J \leq 1000$, can be obtained by taking $\frac{1}{2}$ of the mean range values given in Table X of Tippett (1925) or, for less accurate work, they may be interpolated on a plot of $H(J)$ versus $\log J$, which is nearly linear.

As we have seen in Equation 17, the mean m of g_s exists only for $k > 1$ and the variance of σ^2 only for $k > 2$, and

$$m = \left(\frac{k}{k-1}\right)\mu,$$

$$\sigma^2 = \left(\frac{k}{k-1}\right)^2\left(\frac{k}{k-2}\right)\mu^2.$$

Thus, for $k > 2$, the difference between the signal and noise densities normalized by the standard deviation of the latter is

$$d' = \sqrt{k-2}\left(\frac{\mu}{\nu} - 1\right). \quad [28]$$

If we approach the problem as in TSD, we consider the likelihood ratio

$$L(x) = \left(\frac{\mu}{\nu}\right)^k e^{-(\mu-\nu)x},$$

and so

$$lnL(x) \geq \eta \text{ iff } x \leq c = \frac{kln\frac{\mu}{\nu} - \eta}{\mu - \nu},$$

which, therefore, leads to exactly the same equation as 26, with c replacing $k\beta$.

By substituting $z = \mu/x$ in Equation 26,

$$p(\mathrm{Y}|s) = \int_0^{\mu k\beta} \frac{z^{k-1}e^{-z}}{(k-1)!}dz, \quad [29]$$

which is the incomplete gamma function.

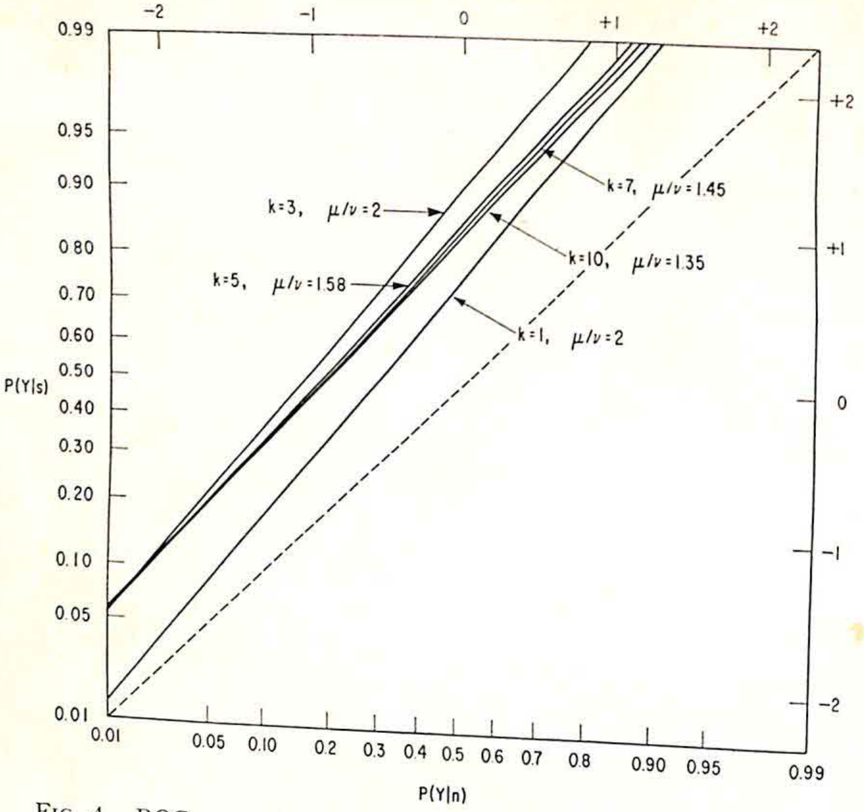


FIG. 4. ROC curve for first κ IATs from J parallel channels. (The parameter k is the product of $J\kappa$. The parameter μ/ν is the ratio of the rate parameter associated with the signal compared to the rate parameter for noise alone.)

For large k, the gamma density approaches the normal $\mathfrak{N}(k,\sqrt{k})$ and so, approximately,

$$p(\mathrm{Y}\,|\,s) \cong \int_{-\infty}^{\sqrt{k}(\mu\beta-1)} \mathfrak{N}(0,1). \qquad [30]$$

A similar expression holds for $p(\mathrm{Y}\,|\,n)$. Let $z(s)$ and $z(n)$ denote the normal deviates corresponding to $p(\mathrm{Y}\,|\,s)$ and $p(\mathrm{Y}\,|\,n)$,

$$z(s) = \frac{\mu}{\nu} z(n) + \sqrt{k}\left(\frac{\mu}{\nu} - 1\right). \qquad [31]$$

Thus, on normal–normal paper, the ROC curve is approximately a straight line with slope μ/ν. Note that by setting $z(n) = 0$,

$$d' = \sqrt{k}\left(\frac{\mu}{\nu} - 1\right), \qquad [32]$$

which differs from the true value by replacing $k - 2$ by k. For measurable levels of detectability ($d' \approx 1$) the slope μ/ν is approximately $1 + 1/\sqrt{k}$ which approaches 1 as k increases. As Figure 4 shows, the normal approximation is excellent for all k; however, the actual slopes are considerably less than μ/ν for k as large as 10. This results from errors in approximating the gamma distribution by the normal distribution.

It should be noted that the shape of the ROC curve is very unusual. For example, with $k = 1$ it is

$$p(\mathrm{Y}\,|\,s) = 1 - [1 - p(\mathrm{Y}\,|\,n)]^{\mu/\nu},$$

which is the power function ROC curve shown in Figure 5. This is different from the power relation,

$$p(\mathrm{Y}\,|\,s) = p(\mathrm{Y}\,|\,n)^{\alpha},$$

suggested by Egan, Greenberg, and Schulman (1961).

Some caution is required in interpreting this prediction since most published data are for fixed-interval designs and, as is shown later (Equation 35), the ROC curve for the corresponding $J = 1$ fixed-interval case is approximately of the correct shape. Moreover, as we have shown empirically (Green & Luce, 1972), even with response-terminated signals, what happens depends on the payoff structure. With an imposed deadline such that responses later than the deadline are penalized, the slope of the ROC curve in a normal probability plot is 1 or slightly less. However, if the deadline is imposed only on signal trials, the slopes are considerably greater than 1.

Speed–Accuracy Tradeoff

At least in situations with response-terminated signals, it has been generally felt that within limits, subjects should be able to increase their accuracy of performance at the expense of taking more time. This is not a matter of variation over the ROC curve, in which case a change in criterion decreases one of the errors, for example, $p(\mathrm{Y} \mid n)$, only at the expense of increasing the other, $p(\mathrm{N} \mid s)$. In a speed–accuracy tradeoff, both types of errors are to be reduced at the expense of taking more time to respond. Empirically, the results have been ambiguous, ranging from fast guess results in which the subject completely fails to identify which signal is presented or identifies it nearly perfectly (Swensson, 1969, Experiment I; Swensson & Edwards, 1971) to difficult-to-discriminate signals in which he exhibits a tradeoff (Swensson, 1969, Experiments II and III).

Within the present model, such a tradeoff can be effected only by using more than one IAT per channel on which to base the response. Obviously, increasing the number of IATs on which to base an estimate of $1/\mu$ decreases the variability of the estimate and takes more time. In fact, from the approximation to $h(J,\kappa)$ above, we see from Equation 23 that

$$\mathrm{MRT}_s(\kappa + 1) - \mathrm{MRT}_s(\kappa)$$
$$\cong \frac{1 + [(\kappa + 2)^{\frac{1}{2}} - (\kappa + 1)^{\frac{1}{2}}]H(J)}{\mu} \geq \frac{1}{\mu}$$

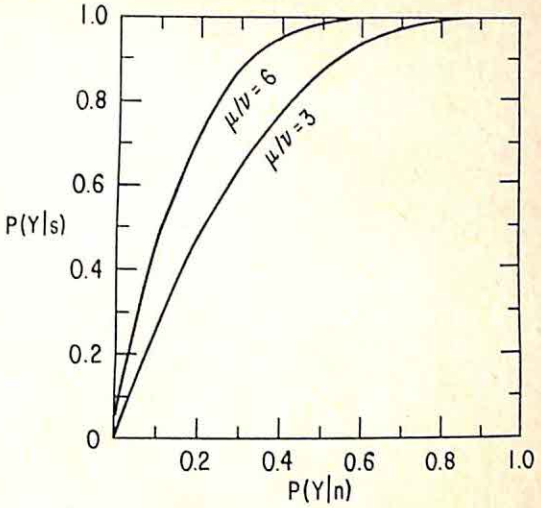


FIG. 5. The $k = 1$ ROC curve for response terminated yes–no design. (The ordinate is the probability of a Y response given that a signal was presented. The abscissa is the probability of a Y response given that no signal was presented. The parameter is the ratio of the rate parameter associated with the signal compared to the rate parameter for noise alone.)

Thus, if the signals are very intense and so μ is large, a negligible change occurs in the response time; whereas, if the signals are very faint and μ is small, a large change occurs and the tradeoff appears to be discontinuous. For example, with $\mu = 4$, the change is at least 250 msec., which is likely to be interpreted as an observer who can operate only in two modes. For signals of moderate intensity, say μ between 20 and 50, we are likely to see a more "continuous" trade of speed and accuracy because in that range each additional IAT adds, on the average, from 20 to 50 msec.

Yes–No Design with Fixed-Signal Duration (J = κ = 1 Model)

The problem of how to deal with interval markers becomes perplexing when we analyze the conventional yes–no design. The situation is exactly the same as the preceding one except that the signal is terminated T seconds after its onset, independent of the subject's response; moreover, the period during which it might be present is clearly marked by a strong signal in another modality. On all trials, we wait for a response to occur, and we record Y to an s presentation as a correct response even if the response occurs long after the termination of the signal.

The major question in analyzing this experimental design is what decision rule to use. The problem is whether or not the nervous system takes into account the information about the end of the signal interval. If it does not and if the first IAT does not lie wholly within the signal interval, then the second pulse is generated by noise, not by signal. If it does, then we have a type of deadline model. We derive the various RT densities and conditional response probabilities for both rules on the assumption $J = 1$, and then show that the second is inconsistent with some RT data. The model for general J is complicated.

Case 1: Decision based on the length of the first IAT independent of when it occurs. This is exactly the same rule as was used above but with $J = 1$; it ignores the information about the end of the signal interval. The equations for n trials are, of course, unchanged (Equation 29 with ν substituted for μ, n for s, and with $k = 1$). The signal case is a good deal more complex because the signal may end before the second pulse arrives. In Appendix B we derive the response density, from which one can derive the following expressions for the response probabilities and MRTs:

$$p(\mathrm{Y}|s) = \begin{cases} 1 - e^{-\mu\beta} - (1 - e^{-\nu\beta})e^{-\mu T} \\ \quad \times \left(\frac{\mu}{\nu} - 1\right), \quad \beta \leq T \\ 1 - e^{-\mu T}e^{-\nu\beta}\left(1 - \frac{\mu}{\nu}\right. \\ \quad \left. + \frac{\mu}{\nu}e^{\nu T}\right), \quad \beta > T \end{cases} \qquad [33]$$

$$p(\mathrm{Y}|n) = 1 - e^{-\nu\beta}. \qquad [34]$$

Eliminating β, we have for the ROC curve

$$p(\mathrm{Y}|s) = \begin{cases} 1 - [1 - p(\mathrm{Y}|n)]^{\mu/\nu} \\ \quad - p(\mathrm{Y}|n)e^{-\mu T}\left(\frac{\mu}{\nu} - 1\right), \\ \quad p(\mathrm{Y}|n) \leq 1 - e^{-\nu T} \\ 1 - [1 - p(\mathrm{Y}|n)]e^{-\mu T} \\ \quad \times \left(1 - \frac{\mu}{\nu} + \frac{\mu}{\nu}e^{\nu T}\right), \\ \quad p(\mathrm{Y}|n) > 1 - e^{-\nu T}. \end{cases} \qquad [35]$$

Observe that the upper limb, $p(\mathrm{Y}|n) > 1 - e^{-\nu T}$, is linear and the lower, curvilinear; some have felt that the data support a curve of this character. For $\beta \leq T$,

$$\begin{aligned} \mathrm{MRT}_{s\mathrm{Y}} = \bar{r}_{\mathrm{Y}} + \Bigg\{ & \frac{2}{\mu}(1 - e^{-\mu\beta}) \\ & - \beta e^{-\mu\beta} + \left(\frac{\mu}{\nu} - 1\right)e^{-\mu T} \\ & \times \left[\beta(1 + e^{-\nu\beta}) - \left(T + \frac{2}{\nu}\right) \right. \\ & \left. \times (1 - e^{-\nu\beta})\right]\Bigg\} \Big/ p(\mathrm{Y}|s) \end{aligned} \qquad [36]$$

$$\begin{aligned} \mathrm{MRT}_{s\mathrm{N}} = \bar{r}_{\mathrm{N}} + \Bigg\{ & \left(\frac{1}{\mu}\right) + \beta\, e^{-\mu\beta} \\ & + \left(\frac{\mu - \nu}{\nu}\right)\left[\left(T + \frac{1}{\nu} + \frac{1}{\mu}\right)\right. \\ & \left. - \left(T + \frac{1}{\nu} + \beta\right)e^{-\nu\beta}\right] \\ & \times e^{-\mu T}\Bigg\} \Big/ p(\mathrm{N}|s). \end{aligned} \qquad [37]$$

The calculation of the MRT equations for $\beta > T$ is left as an exercise.

Some Data

Observe that if the false alarm rate is held constant (i.e., ν and β are constant in Equation 34), then $\mathrm{MRT}_{s\mathrm{Y}}$ (Equation 36) decreases with increasing T, as was reported by Grossberg (1968) for brightness.

In fact, if $\mu\beta$ is sufficiently small so that $1 - e^{-\mu\beta} \cong \mu\beta$, then

$$\mathrm{MRT}_{s\mathrm{Y}} \cong \bar{r}_{\mathrm{Y}} + \begin{cases} \dfrac{1 + \nu\beta - (\mu - \nu)T}{\nu} \\ \qquad \text{if} \quad e^{-\mu T} \cong 1 - \mu T \\ \dfrac{1}{\mu} + \beta, \quad \text{as} \quad T \to \infty. \end{cases}$$

Since Grossberg obtained a nearly negligible false alarm rate, β must have been exceedingly small. So for small T, for example, 10 msec., we see that

$$\mathrm{MRT}_{s\mathrm{Y}} \cong \bar{r}_{\mathrm{Y}} + \begin{cases} \dfrac{1}{\nu}, & \beta, T \ll 1 \\ \dfrac{1}{\mu}, & \mu T \to \infty. \end{cases}$$

Taking as a reasonable guess $\bar{r}_{\mathrm{Y}} = 100$ msec., his data place ν in the range of 3–4 and for these faint signals, μ in the range of 3.5–7.

We next examine whether Bloch's law obtains in our model. For small $\nu\beta$ and νT, Equation 33 reduces to:

$$p(\mathrm{Y}|s) \cong \begin{cases} 1 - e^{-\mu\beta} - (\mu - \nu)\beta e^{-\mu T}, \\ \qquad\qquad \beta \leq T \\ 1 - e^{-\mu T}(1 - \nu\beta + \nu T), \\ \qquad\qquad \beta > T. \end{cases}$$

If we assume both that $\mu \gg \nu$ and μT are sufficiently small so that terms greater than $(\mu T)^2$ can be neglected, then this becomes

$$p(\mathrm{Y}|s) \cong \begin{cases} \beta\mu^2(T - \beta/2), & \beta \leq T \\ \nu\beta(1 - \mu T) + \mu T \\ \qquad - (\mu T)^2/2, & \beta > T. \end{cases}$$

Holding the false alarm rate, and so $\nu\beta$, constant, we see that $p(\mathrm{Y}|s)$ constant implies, for small β, that $\mu^2 T$ is constant, whereas increasing β alters this until, for $\beta > T$, μT is constant.

According to S. S. Stevens (1966), MEs for brief flashes rise as from the .4 to .5 power of I. If the latter, then $\mu^2 T$ constant is equivalent to IT constant, which is Bloch's law. This is well confirmed for small T. According to the model, for $T < \beta$ we should find that it is replaced by the constancy of $I^{\frac{1}{2}}T$, which is not reported. Turning to the loudness data, Stevens and Hall (1966) report that near threshold, $I^{.8}T$ is a constant for equal loudness, which for the model to hold, requires μ to be the .4 power of I. The usually quoted figure is .3; moreover, J. C. Stevens and Hall found that, unlike the brightness data, the exponent is unaffected by the brevity of the signal. However, it is well known that near threshold the apparent exponent is considerably in excess of .3 because the dependence is on $I - I_0$, not on I. So this model does not seem grossly inconsistent with what is known.

Case II: Decision based only on IATs within (*0*,T). The decision rule is this. Choose $\beta \leq T$. Respond Y if at least two pulses occur within $(0, T)$ and if the first IAT is less than β; otherwise, respond N. In Appendix B, we show that

$$p(\mathrm{Y}|s) = 1 - e^{-\mu\beta} - \mu\beta e^{-\mu T}, \quad \beta \leq T. \qquad [38]$$

We cannot get an explicit formula for the ROC curve for $\beta \leq T$ and it is obscure what to assume for $\beta > T$. Numerical calculation yields the former. For the latter, we can either assume the subject stays at the value $\beta = T$, jumps to (1, 1), or forms a linear combination of these two. The latter is shown in Figure 6.

The MRTs (Appendix B) are:

$$\mathrm{MRT}_{s\mathrm{Y}} = \bar{r}_{\mathrm{Y}} + \left[\frac{2}{\mu}(1 - e^{-\mu\beta}) - \beta e^{-\mu\beta} - (\mu T + 1) \times \beta e^{-\mu T}\right] \Big/ p(\mathrm{Y}|s). \qquad [39]$$

$$\mathrm{MRT}_{s\mathrm{N}} = \bar{r}_{\mathrm{N}} + \left[\left(\frac{1}{\mu} + \beta\right) e^{-\mu\beta} + \left(T\mu\beta - \frac{1}{\mu}\right) e^{-\mu T}\right] \Big/ p(\mathrm{N}|s). \qquad [40]$$

This model is almost surely wrong. We note from Equation 39 and $\beta \leq T$, and if $\mu\beta$ is small,

$$\text{MRT}_{s\mathbf{Y}} \to \bar{r}_{\mathbf{Y}} + \begin{cases} \dfrac{\beta}{\mu T} & \text{if} \quad e^{-\mu T} \cong 1 - \mu T \\[2ex] \dfrac{1}{\mu} + \beta & \text{as} \quad T \to \infty. \end{cases}$$

As is easily seen, for $\beta \leq T$, $\dfrac{\beta}{\mu T} < \dfrac{1}{\mu} + \beta$, contrary to Grossberg's (1968) data.

Two-Alternative Forced-Choice Design
(J = κ = 1 Model)

In this well-known design, two listening intervals are defined. They are each of duration T and are separated by a short time. On each trial, the signal is presented in exactly one interval and the subject is required to say which. The possibility exists here, as in the yes–no design, that the subject bases his decision on events outside the listening interval. We assume not, however, and so we assume the natural analogue of the decision rule just studied: the first IAT is observed in each interval and the one having the shorter IAT is reported as containing the signal. The only difficulty, which is serious with weak signals, arises when at least one interval does not contain any IAT; in that case, we let T be the conventional length of the first IAT. Since this can happen with positive probability in each interval, there is a positive probability of an unresolved decision; we assume that with bias b the subject chooses the first interval.

We will not attempt to calculate the densities of decision times, which of course is possible, but will content ourselves with the response probabilities. Let the RV of the time of arrival of the ith pulse be denoted $\mathbf{Y}_i$. Define the decision random variable

$$\mathbf{X} = \begin{cases} \mathbf{Y}_2 - \mathbf{Y}_1, & \text{if} \quad \mathbf{Y}_2 \leq T \\ T, & \text{if} \quad \mathbf{Y}_2 > T. \end{cases}$$

Thus,

$$\begin{aligned} P(\mathbf{X} = t|s) &= \begin{cases} \displaystyle\int_0^{T-t} \mu e^{-\mu x} \mu e^{-\mu t} dx, & t < T \\ \displaystyle e^{-\mu T} + \int_0^T \mu e^{-\mu x} e^{-\mu(T-x)} dx, & t = T \end{cases} \\ &= \begin{cases} \mu(e^{-\mu t} - e^{-\mu T}), & t < T \\ \mu e^{-\mu T}(1 + \mu T), & t = T \end{cases} \end{aligned}$$

$P(\mathbf{X} = t|n)$ is the same with ν substituted for μ. Assuming the independence of the two intervals,

$$\begin{aligned} p(1|s,n) &= P(\mathbf{X}_1 - \mathbf{X}_2 \leq 0) \\ &= \int_0^T P(\mathbf{X}_1 = t|s) \int_t^T P(\mathbf{X}_2 = x|n) dx dt + bP(\mathbf{X}_1 = T|s)\, P(\mathbf{X}_2 = T|n) \\ &= \int_0^T \mu(e^{-\mu t} - e^{-\mu T}) [(e^{-\nu t} + \nu t e^{-\nu T} - (1 + \nu T)e^{-\nu T}] dt + b(1 + \mu T) \\ &\qquad \times (1 + \nu T)\, e^{-(\mu+\nu)T} \\ &= \frac{\mu}{\mu + \nu} + \left(\frac{\nu}{\mu} - 1 - \nu T\right) e^{-\nu T} - \frac{\mu}{\nu} e^{-\mu T} + e^{-(\mu+\nu)T} \left[1 + \frac{\mu}{\nu} - \frac{\nu}{\mu} \right. \\ &\qquad \left. - \frac{\mu}{\mu + \nu} + \mu T + \frac{\mu\nu T^2}{2} + b(1 + \mu T)\,(1 + \nu T) \right]. \qquad [41] \end{aligned}$$

Observe that as $T \to \infty$, the model reduces to the so-called unbiased choice model suggested by various authors (for some history see Luce, 1959, 1963a; see the section on "Discrimination of Intensity" for a generalization).

If we assume that varying b generates the forced-choice ROC curve, then it is given by

$$
\begin{aligned}
p(1 \mid s,n) - p(1 \mid n,s) &= \frac{\mu - \nu}{\mu + \nu} + \left(2\frac{\mu}{\nu} - 1 - \nu T\right)e^{-\nu T} \\
&\quad - \left(2\frac{\nu}{\mu} - 1 - \mu T\right)e^{-\mu T} \\
&\quad + e^{-(\mu+\nu)T}\left[2\frac{\mu}{\nu} - 2\frac{\nu}{\mu} - \frac{\mu - \nu}{\mu + \nu} + (\mu - \nu)T\right]. \quad (42)
\end{aligned}
$$

Note that this is a straight line with slope 1 as has been obtained by similar considerations in the threshold theories (Atkinson, 1963; Krantz, 1969; Luce, 1963a, 1963b; Norman, 1964). Here, in effect, the threshold is not in the subject but in the limited listening interval of the experimental design.

It is important to realize that according to the present theory, the method of generating the ROC curve differs in the yes–no and forced-choice experiments. In the former, a response criterion is varied much as in TSD; in the latter, a response bias is applied to inherently ambiguous information much as in threshold theories. In both cases, the limited period of signal presentation leads to an effective threshold which is reflected in the ROC curves. Eliminating this threshold by having very long intervals ($T \to \infty$) drives the yes–no ROC curve toward a power function and the forced-choice one toward a single point. The present model suggests that different aspects of TSD, threshold, and choice theories will be evident depending on details of the experimental situation.

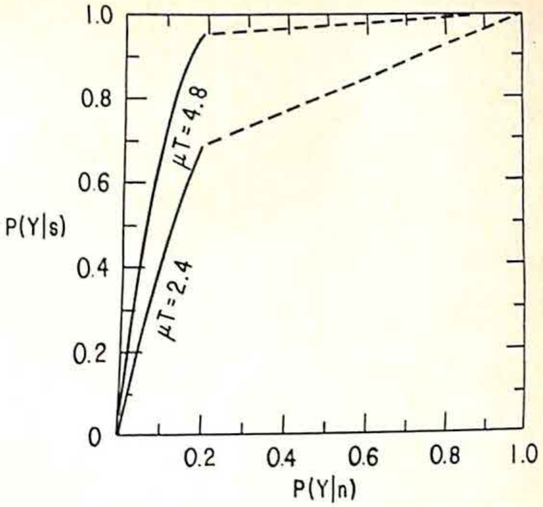


FIG. 6. ROC curve for decision rule based on first IAT within signal interval. (T is the duration of the signal interval, μ the rate parameter associated with the signal. The noise alone rate parameter, ν, times the signal interval was .8 for these calculations, $\nu T = .8$.)

SIMPLE REACTION TIMES WITH RANDOM FOREPERIODS

Consider next a simple RT experiment in which there is a signal on every trial and the subject is asked to respond as soon as he detects it. Unlike the previous designs, there is only one kind of trial and only one response. In order to make the task difficult and to insure that the subject really waits for the signal, it is common experimental practice to make the time from the beginning of the trial (warning signal) to the onset of the reaction signal a random variable. (In our experiments, we have used a discrete approximation to an exponential density because in that case the length of the wait provides absolutely no information about when the signal will occur; some experiments have used other, less desirable distributions.) The only unconventional features of the design we shall study is that the signals are response terminated. This seems appropriate when weak signals are used; it is usually not done with the intense signals used in most RT experiments.

The analysis below was carried out because we had run such an experiment to test a previous theory. It is included here, even though in important aspects it is incomplete, because the experiment is inherently interesting and because, in contrast to the above models, the analysis is mathematically more difficult. The only case that we work out at all fully is $J = \kappa = 1$. To analyze data, however, we really need the $J > 1$ case. The only model we have worked out, which is somewhat different from those above, assumes that the $J = 1$ model applies to each of the channels individually, and that the outputs of these channels all feed into a single channel to which the $J = 1$ model again applies. No work has been done on $\kappa > 1$.

Let $\mathbf{S}$ denote the random time of signal onset after the beginning of the trial, that is, the foreperiod, and let $\mathbf{R}$ denote the time of the response, again measured from the beginning of the trial. Our problem is to derive the distribution of $\mathbf{R}$ under the decision rule that a response is initiated at the end of the first IAT that does not exceed β. The difficulty of the problem lies in the fact that we work not with the first IAT, as in the yes–no designs, but rather with the first IAT that meets a criterion. This difficulty appears to be inherent in any situation where the possible onset of the signal is not specified experimentally.

Analytically, the distribution of $\mathbf{R}$ is composed of its separate distributions at fixed foreperiods, each component distribution being weighted by the probability of that foreperiod. So we first examine the case of a fixed foreperiod, that is, $\mathbf{S} = x$. We, therefore, consider a Poisson process with parameter ν during the interval $[0,x)$ and with parameter μ $(>\nu)$ during the interval $[x,\infty)$. Let $\mathbf{Y}_i$ denote the time of the ith event in this process, where $\mathbf{Y}_{i+1} \geq \mathbf{Y}_i$. Define the decision random variable $\mathbf{D}$ as follows:

$\mathbf{D} = t$ if and only if for some integer j,

$$\mathbf{Y}_j = t, \quad \mathbf{Y}_j - \mathbf{Y}_{j-1} \leq \beta,$$

and for all $i = 1, \cdots, j-2$,

$$\mathbf{Y}_{i+1} - \mathbf{Y}_i > \beta.$$

That is to say, $\mathbf{D}$ is the wait until the end of the first IAT that is $\leq \beta$. We introduce the notation

$$l(x,t) = P(\mathbf{D} = t \mid \mathbf{S} = x) \qquad [43]$$

for the distribution of decision latencies. It is also convenient theoretically to introduce

$$l_0(x,t) = P(\mathbf{D} = t \mid \mathbf{S} = x \text{ and } \mathbf{Y}_1 = 0), \qquad [44]$$

which is the same density subject to the condition that the first event occurs at time 0.

Although most RT data are given in the form of approximations to the following conditional density

$$\begin{aligned} f_{R-S}(t) &= P(\mathbf{R} - \mathbf{S} = t \mid \mathbf{R} \geq \mathbf{S}) \\ &= \int_0^\infty P(\mathbf{R} - \mathbf{S} = t \mid \mathbf{S} = x, \mathbf{R} \geq x) \\ &\qquad \times P(\mathbf{S} = x)\,dx, \end{aligned}$$

it is not sufficient simply to study $l(x,t)$ for $t \geq x$. The reason is that the observed response time density is the convolution of the decision density with the residual one, and so some responses following signal onset arise from decisions reached before it.

Our problem is to derive an equation for $l(x,t)$. This is a rather more difficult task than its simple statement suggests. The only relevant literature which we have found is Gilbert and Pollak (1957) who studied the case $x = 0$; they called it a Poisson coincidence counter. To some degree we will repeat a portion of their work, but it is more difficult than it is worth not to make our treatment of our particular problem self-contained.

Expressions for $l_0(0,t)$

The key to understanding the process is, first, to understand the simplest case in which there is only one Poisson process and the first pulse is at the origin, that is, to solve for l_0. In Appendix C we show that l_0 satisfies

$$l_0(0,t) = \begin{cases} \mu e^{-\mu t}, & 0 \leq t \leq \beta \\ \mu e^{-\mu t} \displaystyle\int_0^{t-\beta} e^{\mu y} l_0(0,y)\,dy, & t > \beta, \end{cases} \qquad [45]$$

which is an integral equation of retarded type with a lag of β. We can solve it in three ways (see Appendix C), each of which may prove useful at one time or another.

First, we have an explicit series expansion; we have found this expansion least useful.

$$l_0(0,t) = \mu e^{-\mu t} \sum_{i=0}^{m-1} a(m-1-i) \times \frac{\mu^i[t-(m-1)\beta]^i}{i!}, \quad (m-1)\beta \leq t < m\beta \qquad [46]$$

where

$$a(0) = 1, \quad a(1) = 0,$$

$$a(m) = \sum_{i=1}^{m-1} \sum_{k=1}^{i} a(i-k)\frac{(\mu\beta)^k}{k!}, \quad \text{for } m > 1. \qquad [47]$$

Second, we work out an approximate solution for small β. The trick is to make a Taylor's series expansion of the integral in Equation 45 about t in powers of β and retain only the linear term. The free constant is determined by the fact that s_0 is a density. This yields the approximation

$$l_0(0,t) \cong \begin{cases} \mu e^{-\mu t}, & 0 \leq t \leq \beta \\ \mu' e^{-\mu'/\mu} e^{-\mu' t} & t > \beta \end{cases} \qquad [48]$$

where

$$\mu' = \frac{\mu^2\beta}{1+\mu\beta}. \qquad [49]$$

And third, we can find the Fourier transform of $l_0(0,t)$; it is surprisingly simple:[10]

$$L_0(0,\omega) = \int_0^\infty e^{-i\omega t} l_0(0,t)dt = \frac{\mu[1-e^{-(i\omega+\mu)\beta}]}{i\omega+\mu-\mu e^{-(i\omega+\mu)\beta}}. \qquad [50]$$

Results for l(x,t) *when* $x \leq \beta$

We now use these expressions for l_0. First, we find the Fourier transform of l (see Appendix C):

$$\begin{aligned} L(x,\omega) &= \int_0^\infty e^{-i\omega t} l(x,t)dt \\ &= \left(\frac{\nu}{i\omega+\nu}\right)^2 \{1-[(i\omega+\nu)x+1] \\ &\quad \times e^{-(i\omega+\nu)x}\} + \frac{e^{(\mu-\nu)x}}{i\omega+\mu}\Big\{L_0(0,\omega) \\ &\quad \times [\nu+(\mu-\nu)e^{-(i\omega+\mu)x}] \\ &\quad + \mu\nu x e^{-(i\omega+\mu)x} - \frac{\mu\nu}{i\omega+\mu} \\ &\quad \times [1-e^{-(i\omega+\mu)x}]\Big\}, \end{aligned} \qquad [51]$$

where $L_0(0,\omega)$ is given in Equation 50.

Assuming for the moment that we have estimates of the parameters ν, μ, and β, we can use Equation 51 as follows. With a fixed foreperiod, the response density is a convolution of the decision time with the residual time, so its Fourier transform $F(x,\omega)$ can be written

$$F(x,\omega) = L(x,\omega)R(\omega).$$

Or, to use more data, if we have a distribution g of foreperiods,

$$F(\omega) = \int_0^\beta g(x)F(x,\omega)dx = \int_0^\beta g(x)L(x,\omega)dxR(\omega).$$

This permits us to calculate R and so r from the data using Equation 51. The only difficulty is in estimating the parameters, with which we deal below.

For $x \leq \beta$, we can compute the MRT from a signal onset at time x by using Equation 51 and the fact that

$$\int_0^\infty tf(t)dt = i\lim_{\omega\to 0}\frac{dF(\omega)}{d\omega}.$$

---

[10] As J. W. Tukey suggested to us, a simple direct argument leads to Equation 50. The process consists of $n=1, 2, \cdots$, independent drawings from the exponential distribution, the first $n-1$ of which are $>\beta$ and the nth is $\leq\beta$. As is easily computed, these two truncated densities have transforms

$$U(\omega) = \frac{\mu e^{-i\omega\beta}}{\mu+i\omega} \quad \text{and} \quad V(\omega) = \frac{\mu[1-e^{-(\mu+i\omega)\beta}]}{(1-e^{-\mu\beta})(\mu+i\omega)}$$

and probabilities of occuring $e^{-\mu\beta}$ and $1-e^{-\mu\beta}$, respectively. Thus,

$$L_0(0,\omega) = \sum_{n=1}^{\infty}[e^{-\mu\beta}U(\omega)]^{n-1}(1-e^{-\mu\beta})V(\omega) = \frac{(1-e^{-\mu\beta})V(\omega)}{1-e^{-\mu\beta}U(\omega)},$$

which is Equation 50.

Carrying out that computation,

$$\begin{aligned}\mathrm{MRT}_x &= \bar{r} + i\lim_{\omega\to 0}\frac{dL(x,\omega)}{d\omega} - x\\ &= \bar{r} + \frac{2}{\nu}(1 - e^{-\nu x}) + (A - Bx)e^{-\nu x}\\ &\quad + Ce^{(\mu-\nu)x} - x, \quad x \leq \beta, \quad [52]\end{aligned}$$

where

$$A = \left(\frac{\mu - \nu}{\mu^2}\right)\left(\frac{1}{1 - e^{-\mu\beta}}\right) + \frac{1}{\mu} + \frac{\nu}{\mu^2}$$

$$B = 1 - \frac{\nu}{\mu}$$

$$C = \frac{\nu}{\mu^2}\left(\frac{1}{1 - e^{-\mu\beta}} - 1\right).$$

For $\nu = 0$ or $x = 0$,

$$\mathrm{MRT}_x = \bar{r} + \frac{1}{\mu}\left(\frac{1}{1 - e^{-\mu\beta}} + 1\right) \quad [53]$$

and

$$\left(\frac{\partial \mathrm{MRT}_x}{\partial x}\right)_{x=0} = -\frac{\nu}{\mu} < 0.$$

It is evident that as the mean foreperiod is increased, the subject can maintain a fixed false alarm rate only by decreasing β. Since this increases both A and C, MRT_x, $x \leq \beta$, must increase with increasing mean foreperiod. This was observed in Green and Luce (1971). Unfortunately, we cannot check their observations in greater detail until we know more about $l(x,t)$ for $x > \beta$.

MRT Data

Assuming that the noise is negligible, so Equation 53 applies, and recalling from Equation 17 that the mean (or geometric mean) MEs are proportional to μ, that is, $\mathrm{ME} = D\mu$, then

$$\mathrm{MRT} \cong \bar{r} + \begin{cases}\dfrac{2D}{\mathrm{ME}}, & \text{for } \mu \text{ large}\\[2ex] \dfrac{D}{\mathrm{ME}} + \dfrac{D^2}{\mathrm{ME}^2\beta}, & \text{for } \mu \text{ small}\\ & (1 - e^{-\mu\beta} \cong \mu\beta).\end{cases}$$

Vaughan, Costa, and Gilden (1966) showed that the data of several authors for visual signals of fixed duration are consistent with the formula

$$\mathrm{MRT} = \bar{r} + D'/I^{1/3},$$

where I is luminance. As they point out, it is noteworthy that "subjective brightness," as determined by the method of ME, grows as $I^{1/3}$ (J. C. Stevens & Stevens, 1963; S. S. Stevens, 1957; S. S. Stevens & Galanter, 1957). Assuming that the magnitude estimates are proportional to $I^{1/3}$, then at least for moderately large μ we would expect Vaughan et al.'s result. The only complication in this argument is that they obtained the data using exceedingly brief signals—in their own data, $T = 10$ μsec. (in those of others $T = 55$ msec. to 575 msec.). For these data to be compatible with our theory, it is necessary to suppose that the effect of a very brief visual signal on the neural pulse train is much longer than the signal's duration.

For loudness, McGill (1960) has interpreted his data as showing MRT is linear in $\mathrm{ME}^{-1/3}$. Taking the data reported in McGill (1963), unpublished data of J. G. Snodgrass (personal communication, 1962), and those of Chocholle (1940) and choosing $\bar{r}$ to be 5 msec. smaller than the MRT to the loudest signal, we obtain the curves shown in Figure 7. Since Chocholle did not obtain MEs, we have plotted his data as if the magnitude exponent were .3 (his other two subjects give nearly identical results). These plots seem to support the theory; however, whether or not one says the slopes are -1, as predicted, depends on exactly what choice one makes for $\bar{r}$. Altering it by only a few milliseconds has an enormous effect on the apparent slope. McGill made a different choice from ours. Chocholle's data, which cover the range down to the subject's threshold, exhibit the predicted steepening of slope at low intensities.

Estimation of Parameters

We need schemes to estimate ν, μ, and β. As we shall see, we have satisfactory methods for estimating ν and β, but not for μ except when β is small. For other β we do not presently have a method for estimating μ.

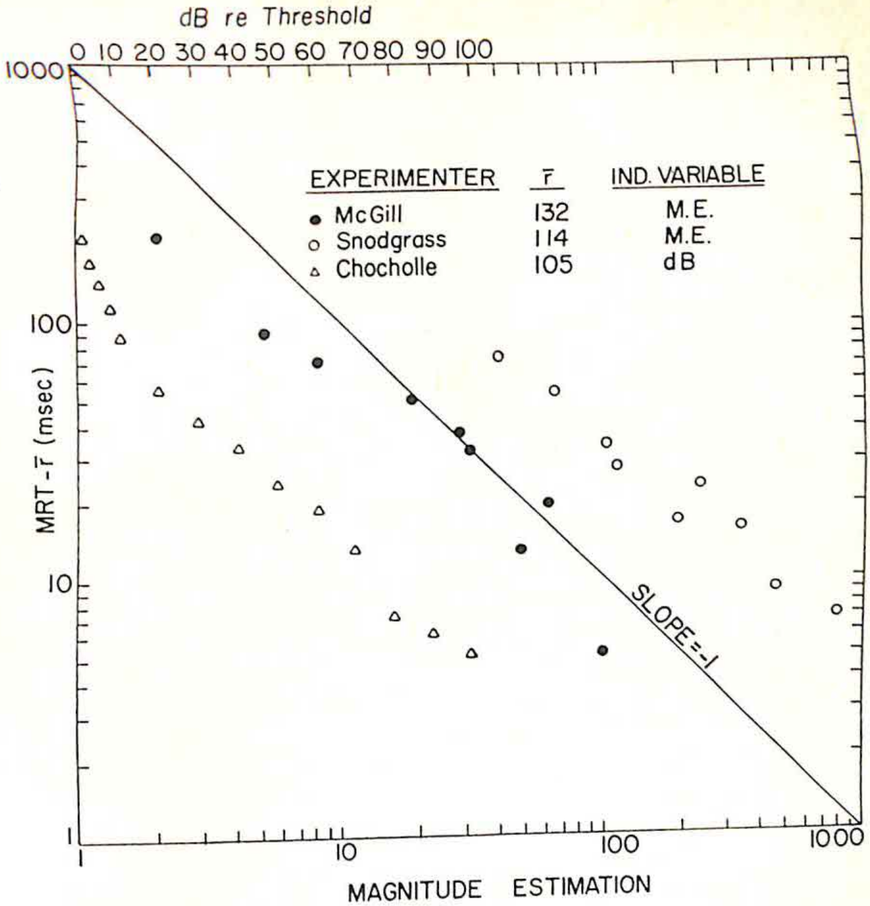


FIG. 7. Plots of MRT $-\ \bar{r}$ versus ME of loudness of 1000-Hz. tones from three experiments. (For Chocholle's data, it is assumed that ME $= \alpha I^{.3}$. In each case, $\bar{r}$ = MRT$_{\min}$ − 5 [in milliseconds]; for the Snodgrass data, the loudest signal [ME = 1500] is not on the graph.)

On the assumption of a known small β, then to estimate μ it is sufficient to estimate μ' and use Equation 49. To estimate μ' we use the approximate expression for $l_0(0,t)$, Equation 48 (which is where the assumption of small β enters), in equations for $l(x,t)$, convolute the result with r, and integrate that over an arbitrary distribution of foreperiods. As we show in Appendix C, the resulting density of reaction times has tails of the following form:

$$f_{R-S}(t) = P(\mathrm{R} - \mathrm{S} = t \mid \mathrm{R} \geqslant \mathrm{S})$$
$$\cong Ae^{-\mu t} + Be^{-\mu' t},$$
$$t > \tau + 2\beta, \quad [54]$$

where μ' is given by Equation 49 and A and B depend on ν, μ, r, and the distribution of foreperiods. Observe that

$$\frac{\mu'}{\mu} < \frac{\mu\beta}{1 + \mu\beta},$$

and so for β sufficiently small, μ' is appreciably smaller than μ. Thus, $e^{-\mu' t}$ is the dominant term in Equation 54 and so the tail of the distribution of RTs allows us to estimate μ'.

To get at the parameter ν, consider $l(x,t)$ for $t \leq x$. It can be shown (Appendix C) that $l(x,t) = l(t,t)$ for $t \leq x$ and that it satisfies the following integral equation of retarded type:

$$l(t,t) = \begin{cases} \nu^2 t e^{-\nu t}, & t \leq \beta \\ \nu e^{-\nu t}\left(\displaystyle\int_0^{t-\beta} e^{\nu z} l(z,z)\,dz + \nu\beta\right), & \\ & t > \beta. \quad [55] \end{cases}$$

To calculate the Fourier transform, we note that $l(t,t) = l(\infty,t)$, and the latter is simply the convolution of $\nu e^{-\nu t}$ with $l_0(\infty,t)$. Therefore,

$$L(\omega) = L(\infty,\omega) = \frac{\nu}{\nu + i\omega} L_0(\infty,\omega).$$

However, $L_0(\infty,\omega)$ is just $L_0(0,\omega)$ with μ replaced by ν, so by Equation 50,

$$L(\omega) = \frac{\left(\dfrac{\nu}{i\omega + \nu}\right)^2 [1 - e^{-(i\omega+\nu)\beta}]}{1 - \dfrac{\nu}{i\omega + \nu} e^{-(i\omega+\nu)\beta}}. \qquad [56]$$

Moreover, if we assume that for $t > \tau'$

$$l(t,t) \cong Ce^{-\nu' t}, \qquad [57]$$

then direct substitution into Equation 55, collecting terms in $e^{-\nu t}$ and $e^{-\nu' t}$, and choosing ν' and C so that the equation is satisfied yields

$$\nu' = \nu[1 - e^{-(\nu-\nu')\beta}], \quad C = \nu\beta(\nu - \nu'). \qquad [58]$$

Let f_R denote the density of response times for which the response precedes the signal onset. If we assume that the density of foreperiods is $\lambda\, e^{-\lambda t}$, then

$$f_R(t) = Ke^{-(\lambda+\nu')t}, \quad t \geq \tau' + \tau, \qquad [59]$$

where K depends on ν, β, and r. Thus, we can estimate ν' from the tail of f_R.

In order to estimate ν and β, we need another expression relating them. We turn to that next.

False Alarm Probability

If we use an exponential distribution of foreperiods, the probability of a false alarm is easily seen to be given by

$$P(\mathrm{R} < \mathrm{S}) = \int_0^\infty e^{-\lambda x} \int_0^x l(y,y) r(x - y)\,dy\,dx,$$

which is simply the LaPlace transform of the convolution of l and r, and so it is the product of their respective LaPlace transforms. By the same technique as used to get the Fourier transform of $l(t,t)$ in Equation 56, we obtain

$$P(\mathrm{R} < \mathrm{S}) = \frac{\left(\dfrac{\nu}{\lambda + \nu}\right)^2 [1 - e^{-(\lambda+\nu)\beta}]}{1 - \dfrac{\nu}{\lambda + \nu} e^{-(\lambda+\nu)\beta}} R_{-\lambda}, \qquad [60]$$

where

$$R_{-\lambda} = \int_0^\infty e^{-\lambda t} r(t)\,dt. \qquad [61]$$

Observe that by solving Equation 58 for β and substituting it in Equation 60, we obtain

$$\frac{P(\mathrm{R} < \mathrm{S})}{R_{-\lambda}} = \frac{1 - f(x,\alpha)}{(1 + \alpha x)[1 + \alpha x - f(x,\alpha)]},$$

where

$$x = \nu'/\nu$$
$$\alpha = \lambda/\nu'$$
$$f(x,\alpha) = (1 - x)^{(1+\alpha x)/(1-x)}.$$

This function is plotted in Figure 8. It can be used as follows: Since λ is determined by the experimenter and ν' can be estimated by Equation 58, α is known. Observe that for λ sufficiently small,

$$R_{-\lambda} \cong 1 - \lambda\bar{r}.$$

We can determine $\bar{r}$ from a very intense signal, so $P(\mathrm{R} < \mathrm{S})/R_{-\lambda}$ is also known. Thus, x is determined, from which we obtain ν, and so β is determined from the plot of $\beta\nu$ (Equation 58) also included in Figure 8.

Moments of the Response Time Distribution

If the signal onset is exponentially distributed, the overall response time density is

$$\int_0^\infty \lambda e^{-\lambda x} l(x,t)\,dx.$$

Its Fourier transform is

$$L_\lambda(\omega) = \int_0^\infty e^{-i\omega t} \int_0^\infty \lambda e^{-\lambda x} l(x,t)\,dx\,dt$$
$$= \int_0^\infty \lambda e^{-\lambda x} L(x,\omega)\,dx,$$

where L is given by Equation 51 for $x \leq \beta$. If an explicit expression could be obtained

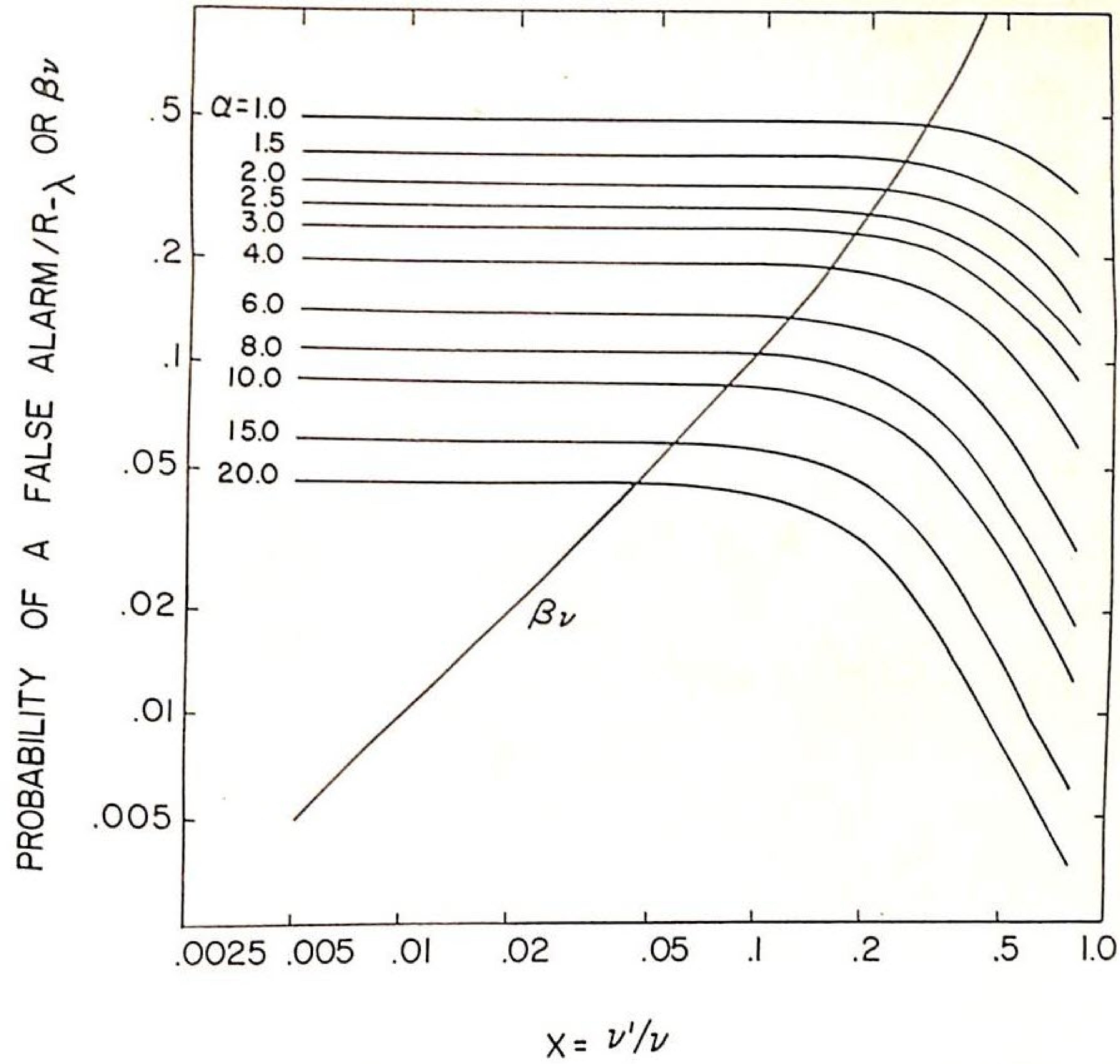


Fig. 8. Plots of $P(\mathbf{R} < \mathbf{S})/R_{-\lambda}$ versus $x = \nu'/\nu$ with $\alpha = \lambda/\nu'$ as a parameter and of $\beta\nu$ versus x. (Assuming $P(\mathbf{R} \leq \mathbf{S})/R_{-\lambda}$ and α known, this plot may be used to estimate ν and β.)

for $x > \beta$, then the moments could be obtained by taking derivatives of $L_\lambda(\omega)$ and letting $\omega \to 0$. For example, the mean would be given by

$$\lim_{\omega \to 0} i \frac{dL_\lambda(\omega)}{d\omega}.$$

In principle, this gives us all the additional equations needed to estimate the parameters without invoking any approximations.

Estimates of μ for Sound Intensity

It is easy to estimate $\mu_{\max}/\mu_{\min}$ from the empirical fact that, roughly, $I_{\max}/I_{\min}$ is 10^{10}. So, from the ME exponent of .3, we conclude $\mu_{\max}/\mu_{\min}$ is about 10^3. It should be realized that this estimate is considerably larger than the observed range of pulse rates in the peripheral fibers (Grossman & Vierstein, 1961; Kiang, 1965; Rodieck et al., 1962). A range of 10^2 is unusual; 10 is much more common. This suggests either some sort of transfer from one neural path to another as intensity increases or that the central system has considerably more dynamic range.

We do not know of psychophysical data that permit us to estimate either $\mu_{\min}$ or $\mu_{\max}$. The nearest that we can come to it, which illustrates the method we would use to estimate $\mu_{\min}$ if the data were available, is to estimate the value ν, corresponding to the noise background used in our experiments, namely, a noise power density of 40 db. We use two methods. The first probably underestimates ν because it assumes $J = 1$. It is based on the analysis of the random foreperiod RT experiment. We estimate ν and β by the method suggested following Equation 60. According to Green and Luce (1971), in an experiment with mean exponential foreperiod of 2 seconds ($\lambda = .5$) and with 10,000 observations collected from each of two subjects, the tail of f_R (Equation 59) is exponential, $\nu' \cong .07$, and $P(\mathbf{R} < \mathbf{S}) \cong .1$.

If we assume, as is plausible, $\bar{\tau} = .2$, then $R_{-\lambda} = 1 - (.5)(.2) = .9$. Solving in Figure 8, $\nu \cong .35$ and $\beta \cong .8$.

The second method is based on fixed-interval yes–no data and on the assumption that the response rule ignores the termination of the interval. In that case, the noise equation is the same as in the response terminated case, Equation 26. With a reasonably strict criterion, it is probably safe to assume $\kappa = 1$, $\beta \leq .2$, and that $p(Y|n) \cong .1$. If so, using the normal approximation to the gamma, we see by Equation 30

$$J^{\frac{1}{2}}(\nu\beta - 1) = z(n) \cong -1.30.$$

Using our estimate of $J \cong 27$, we obtain $\nu \geq 3.72$. Note that this corresponds well with the estimate from Grossberg's (1968) MRT data for brightness.

With the latter estimate, a signal 65 db. above the noise would have a value of μ/ν of about $10^{6.5(.3)} = 90$, so $\mu \geq 335$, For $T = \frac{1}{2}$ second, $\mu T \geq 167 > 65 \cong K$, hence these numbers are consistent with our definition of a strong signal.

According to the data of Rose et al. (1967), the representation of signals of frequencies not greater than 2500 Hz. in the peripheral neurons is quantized in the period of the signal. Thus, a maximum rate for a 1000-Hz. signal would be a pulse every millisecond, that is, a value of 1000 for μ. This does not appear grossly inconsistent with the psychophysical estimates.

The order of magnitude difference in our estimates of ν, .35 and 3.72, probably results from the fact that we have incorrectly assumed $J = 1$ in analyzing the random foreperiod experiment. To show that this may account for the discrepancy, consider the following $J > 1$ model for this experiment. Each of the J parallel channels processes pulses as we have assumed; however, whenever two successive pulses are less than δ apart, it triggers not a response, but rather a pulse on a common channel. If δ is small, say a few milliseconds, this could be done by energy summation at a synapse with a threshold. The common channel, then, carries the superposition of the outputs of the J parallel channels. When there is no signal, the pulses on the common channel should not cluster; when there is, they should cluster. Moreover, if J is sufficiently large, the process on the common channel is approximately Poisson, even though the recurrent events on the parallel channels are not (see Cox, 1962, pp. 71–79). Finally, we assume that the common channel triggers a response according to the $J = 1$ model with a criterion $\beta \neq \delta$.

From Equation 50, we obtain

$$1/\nu(1 - e^{-\nu\delta})$$

as the mean time between successive outputs of each of the parallel channels. Thus, the intensity parameter of the common channel is

$$\nu' = J\nu(1 - e^{-\nu\delta}) \cong J\nu^2\delta.$$

So, if $\nu = 3.72$, $J = 27$, and $\delta = .001$, then $\nu' = .373$, which seems to account for the order of magnitude difference in our estimate.

According to the arguments leading to Equations 54 and 59, the tails of both the false alarm and RT densities should be exponential in a simple RT experiment with an exponential distribution of foreperiods. Luce and Green (1970) verified this prediction, and they estimated[11] the parameters μ' and ν' for various signal intensities. Let μ_1 and ν_1 denote the corresponding true single-channel values which are related to μ' and ν' by Equations 49 and 58, respectively. Finally, let μ and ν be the underlying parameters on the several parallel channels, as just derived. If $\mu_1\beta$ is small

[11] The cut point used in analyzing these data was 500 msec. Recall that Equation 54 was derived for $t > \tau + 2\beta$. We are probably safe in assuming τ is not greater than 300 msec.; this is based on simple RT distributions to intense signals, which according to this theory is dominated by τ. So the above inequality would hold if β were less than 100 msec. However, using the method for estimating β outlined in connection with Figure 8, we estimate β for these data to be about 1 second. Therefore, we have no reason to believe our approximations are applicable. In spite of that, we continue the calculation as if they are justified because the results seem consistent with earlier ones.

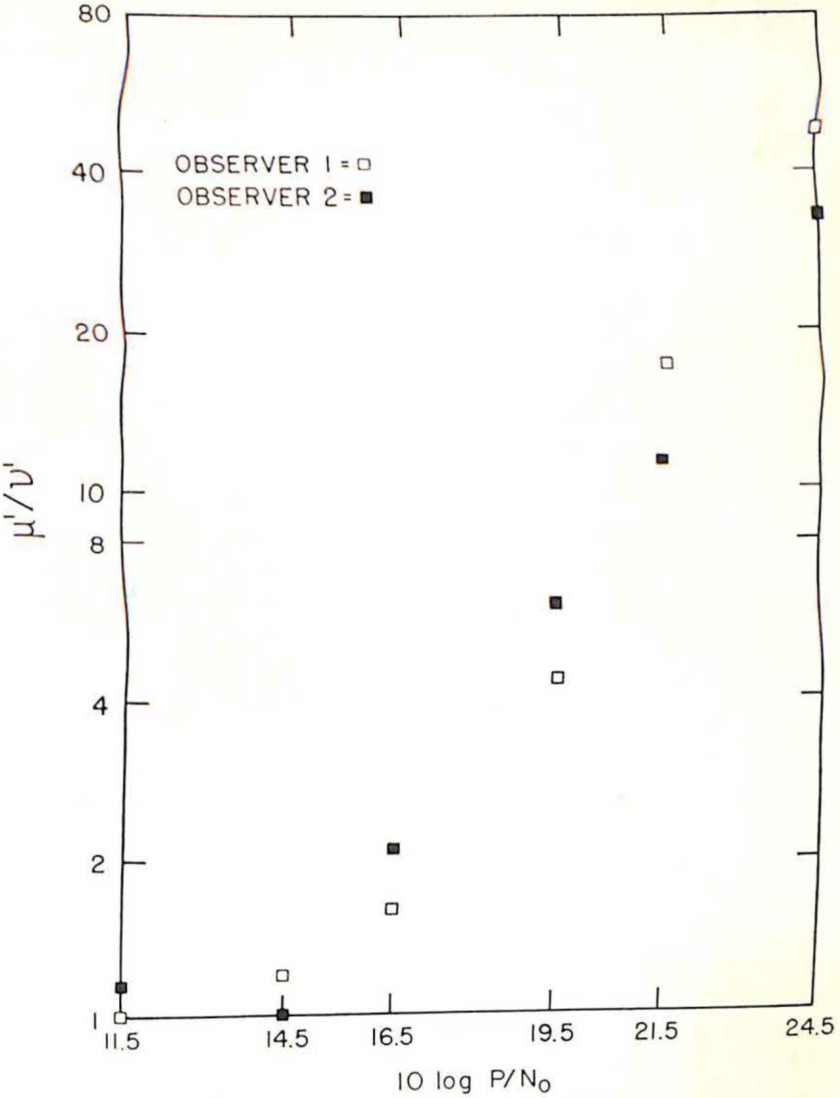


FIG. 9. Estimated μ'/ν' versus the ratio of signal power to noise density in decibels. (The two Poisson parameters were estimated from the tails [> 500 msec.] of RT densities. This is Figure 9 of Luce & Green, 1970.) (Reprinted from an article by R. D. Luce and D. M. Green in *Perception and Psychophysics*, Vol. 7. Copyrighted by Psychonomic Journals, Inc., 1970.)

relative to 1, then

$$\frac{\mu'}{\nu'} \cong \frac{\mu_2{}^2\beta}{1+\mu_1\beta} \cdot \frac{1+\nu_1\beta}{\nu_2{}^2\beta} \cong \left(\frac{\mu_1}{\nu_1}\right)^2 \cong \left(\frac{\mu}{\nu}\right)^4.$$

Figure 9 presents the empirical plot of μ'/ν' versus signal-to-noise in decibels for two subjects. We see that to a good approximation μ'/ν', and so by what we just showed μ/ν, is a power function of $I(s)/I(n)$, confirming again one of our basic postulates. Moreover, taking into account the factor of 4, the estimated exponent γ is .4, which is well within the range of values found by magnitude estimation of the intensity of pure tones.

For a particular signal-to-noise ratio, we can also estimate μ/ν from response terminated yes–no ROC data. For example, we have found that 10 log P/N_0 = 16.5 db. yields a value of about 1 for d', that is, $z(s) = 1$ and $z(n) = 0$. From Equation 32, we see that

$$\frac{\mu}{\nu} = 1 + 1/J^{\frac{1}{2}},$$

which with our estimate of $J \cong 27$ yields $\mu/\nu = 1.193$. By the above argument, the observed value of μ'/ν' should be $(1.193)^4 = 1.85$, which falls between the two data points. If we reverse the argument, this provides an independent estimate of J which agrees with the earlier one from ME data.

SUMMARY

The postulates of the theory are: signal intensity is transformed into J identical Poisson processes on parallel channels; the (intensity) parameter of these Poisson processes is a power function, with exponent γ, of signal intensity; decisions are based on the IATs of pulses, which are estimates of the Poisson parameter; up to K IATs per channel can be collected and stored prior to reaching a decision; and observed responses are delayed beyond the decision time by the addition of a residual latency which is no longer than τ seconds.

These ideas were worked out in some detail for various standard types of experiments, and a number of well-known results were shown to follow, several postulates were tested, and estimates were made of the several parameters.

Aside from neurophysiological data on peripheral nerves, the strongest evidence for the Poisson assumption comes from the exponential decay of RT densities. Support for the power function came from three direct sources—ME data, the nature of the decrease in MRT with signal intensity, and a plot of Poisson parameters (estimated from the tails of RT densities) versus signal intensity—and one indirect one—the prediction of Weber's law. It is gratifying that all three estimates of the exponent γ were about the same, .3–.4 for sound intensity. Actual values of the Poisson parameters, based on yes–no and RT data, suggest a value of ν, corresponding to the noise background, of about 3–4, or a mean IAT of 250–350 msec. A signal-to-noise ratio corresponding to a d' of 1 yields a ratio of μ/ν of about 1.2.

The estimates of the constants J and K are less secure, and future work should attempt to find additional ways to estimate them. From Weber's fraction, we concluded that for 1000-Hz. pure tones, JK has a value of about 1750. Two arguments, one based on the ratio of the mean to standard deviation of MEs and another on a yes–no ROC curve and a previous estimate of μ/ν, suggested that J is about 27, in which case K is about 65. Finally, we estimated τ to be about 300 msec. on the ground that simple RTs to intense signals are very nearly the residual latency.

REFERENCES

ATKINSON, R. C. A variable sensitivity theory of signal detection. *Psychological Review*, 1963, **70**, 91–106.

BARLOW, H. B. Retinal noise and absolute threshold. *Journal of the Optical Society of America*, 1956, **46**, 634–639.

BARLOW, H. B. Increment thresholds at low intensities considered as signal/noise discriminations. *Journal of Physiology*, 1957, **136**, 469–488.

BATHER, J. A. On a quickest detection problem. *Annals of Mathematical Statistics*, 1967, **38**, 711–724.

BELLMAN, R., & COOKE, K. L. *Differential-difference equations*. New York: Academic Press, 1963.

BOUMAN, M. A. History and present status of quantum theory in vision. In W. A. Rosenblith (Ed.), *Sensory communication*. Cambridge: M.I.T. Press & Wiley, 1961.

BRAIDA, L. D., & DURLACH, N. I. Intensity perception, II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, 1972, **52**, in press.

BRUGGE, J. F., ANDERSON, D. J., HIND, J. E., & ROSE, J. E. Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds. *Journal of Neurophysiology*, 1969, **32**, 386–401.

CARTERETTE, E. C. Random walk models for reaction times in signal detection and recognition. *Proceedings of the XVIII International Congress of Psychology*, 1966, **16**, 84–95.

CHOCHOLLE, R. Variations des temps de réaction auditifs en fonction de l'intensité à diverses fréquences. *L'Année Psychologique*, 1940, **41**, 5–124.

CHRISTIE, L. S., & LUCE, R. D. Decision structure and time relations in simple choice behavior. *Bulletin of Mathematical Biophysics*, 1956, **18**, 89–112.

COX, D. R. *Renewal theory*. London: Metheun, 1962.

CREELMAN, C. D. Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, 1961, **34**, 582–593. Reprinted in J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1964.

DURLACH, N. I., & BRAIDA, L. D. Intensity perception, I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, 1969, **46**, 372–383.

Edwards, W. Optimal strategies for seeking information: Models for statistics, choice reaction times and human information processing. *Journal of Mathematical Psychology*, 1965, **2**, 312–329.

Egan, J. P., Greenberg, G. Z., & Schulman, A. I. Operating characteristics, signal detectability, and the method of free response. *Journal of the Acoustical Society of America*, 1961, **33**, 993–1007.

El'sgol'ts, L. E. *Introduction to the theory of differential equations with deviating arguments.* San Francisco: Holden-Day, 1966.

Estes, W. K. A random walk model for choice behavior. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford: Stanford University Press, 1960.

Gilbert, E. N., & Pollak, H. O. Coincidences in Poisson patterns. *Bell System Technical Journal*, 1957, **36**, 1005–1033.

Green, D. M., & Luce, R. D. Detection of auditory signals presented at random times. *Perception and Psychophysics*, 1967, **2**, 441–449.

Green, D. M., & Luce, R. D. Detection of auditory signals presented at random times, III. *Perception and Psychophysics*, 1971, **9**, 257–268.

Green, D. M., & Luce, R. D. Speed-accuracy tradeoff in auditory detection. In S. Kornblum (Ed.), *Attention and performance.* Vol. 4. New York: Academic Press, 1972, in press.

Green, D. M., & McGill, W. J. On the equivalence of detection probabilities and well-known statistical quantities. *Psychological Review*, 1970, **77**, 294–301.

Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966.

Grice, G. R. Stimulus intensity and response evocation. *Psychological Review*, 1968, **75**, 359–373.

Grossberg, M. The latency of response in relation to Bloch's law at threshold. *Perception and Psychophysics*, 1968, **4**, 229–232.

Grossman, R. G., & Vierstein, L. J. Discharge patterns of neurons in cochlear nucleus. *Science*, 1961, **134**, 99.

Hind, J. E., Anderson D. J., Brugge, J. F., & Rose, J. E. Coding of information pertaining to paired low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, 1967, **30**, 794–816.

John, I. D. A statistical decision theory of simple reaction time. *Australian Journal of Psychology*, 1967, **19**, 27–34.

Kendall, M. G., & Stuart, A. *The advanced theory of statistics: I.* New York: Hafner, 1958.

Kiang, N. Y-S. *Discharge patterns of single fibers in the cat's auditory nerve.* Cambridge: M.I.T. Press, 1965.

Kiang, N. Y-S. A survey of recent developments in the study of auditory physiology. *Annals of Otology, Rhinology and Laryngology*, 1968, **77**, 656–676.

Kinchla, R. A., & Smyzer, F. A diffusion model of perceptual memory. *Perception and Psychophysics*, 1967, **2**, 219–229.

Kohfeld, D. L. Effects of the intensity of auditory and visual ready signals on simple reaction time. *Journal of Experimental Psychology*, 1969, **82**, 88–95.

Krantz, D. H. Threshold theories of signal detection. *Psychological Review*, 1969, **76**, 308–324.

Kristofferson, A. B. Successiveness discrimination as a two-state quantal process. *Science*, 1967, **158**, 1337–1339. (a)

Kristofferson, A. B. Attention and psychophysical time. *Acta Psychologica*, 1967, **27**, 93–100. (b)

LaBerge, D. L. A recruitment theory of simple behavior. *Psychometrika*, 1962, **27**, 375–396.

Laming, D. R. J. *Information theory of choice-reaction times.* New York: Academic Press, 1968.

Luce, R. D. *Individual choice behavior.* New York: Wiley, 1959.

Luce, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963. (a)

Luce, R. D. A threshold theory for simple detection experiments. *Psychological Review*, 1963, **70**, 61–79. (b)

Luce, R. D. A model for detection in temporally unstructured experiments with a Poisson distribution of signal presentations. *Journal of Mathematical Psychology*, 1966, **3**, 48–64.

Luce, R. D., & Green, D. M. Detection of auditory signals presented at random times, II. *Perception and Psychophysics*, 1970, **7**, 1–14.

Luce, R. D., & Mo, S. S. Magnitude estimation of heaviness and loudness by individual subjects: A test of a probabilistic response theory. *British Journal of Mathematical and Statistical Psychology*, 1965, **18**, 159–174.

McGill, W. J. Loudness and reaction time: A guided tour of the listener's private world. *Proceedings of the XVI International Congress of Psychology, 1960.* Amsterdam: North-Holland, 1960.

McGill, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963.

McGill, W. J. Neural counting mechanisms and energy detection in audition. *Journal of Mathematical Psychology*, 1967, **4**, 351–376.

McGill, W. J., & Goldberg, J. P. Pure-tone intensity discrimination and energy detection. *Journal of the Acoustical Society of America*, 1968, **44**, 576–581.

Norman, D. A. Sensory thresholds, response biases, and the neural quantum theory. *Journal of Mathematical Psychology*, 1964, **1**, 88–120.

Parzen, E. *Stochastic processes.* San Francisco: Holden-Day, 1962.

Perkel, D. H., & Bullock, T. H. Neural coding. *Bulletin of the Neurosciences Research Program*, 1968, **6**, 221–348.

Rodieck, R. W., Kiang, N. Y-S., & Gerstein, G. L. Some quantitative methods for the study of spontaneous activity of single neurons. *Biophysics Journal*, 1962, **2**, 351–368.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, 1967, **30**, 769–793.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. Some possible neural correlates of combination tones. *Journal of Neurophysiology*, 1969, **32**, 402–423.

Schneider, B., & Lane, H. Ratio scales, category scales, and variability in the production of loudness and softness. *Journal of the Acoustical Society of America*, 1963, **35**, 1953–1961.

Shiryaev, A. N. On optimum methods in quickest detection problems. *Theory of Probability and Its Applications*, 1963, **8**, 22–46.

Siebert, W. M. Some implications of the stochastic behavior of primary auditory neurons. *Kybernetik*, 1965, **2**, 206–215.

Siebert, W. M. Stimulus transformations in the peripheral auditory system. In P. A. Kohlers & M. Eden (Eds.), *Recognizing patterns*. Cambridge, Mass.: M.I.T. Press, 1968.

Siebert, W. M. Frequency discrimination in the auditory system: Place or periodicity mechanisms? *Proceedings of the IEEE*, 1970, **58**, 723–730.

Stevens, J. C., & Hall, J. W. Brightness and loudness as functions of stimulus duration. *Perception and Psychophysics*, 1966, **1**, 319–327.

Stevens, J. C., & Stevens, S. S. Brightness function: Effects of adaptation. *Journal of the Optical Society of America*, 1963, **53**, 375–385.

Stevens, S. S. On the psychophysical law. *Psychological Review*, 1957, **64**, 153–181.

Stevens, S. S. Duration, luminance, and the brightness exponent. *Perception and Psychophysics*, 1966, **1**, 96–100.

Stevens, S. S. Neural events and the psychophysical law. *Science*, 1970, **170**, 1043–1050.

Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 1957, **54**, 377–411.

Stone, M. Models for choice-reaction time. *Psychometrika*, 1960, **25**, 251–260.

Swensson, R. G. The elusive tradeoff: Speed versus accuracy in a visual discrimination task. Unpublished doctoral dissertation, University of Michigan, 1969.

Swensson, R. G., & Edwards, W. Response strategies in a two-choice reaction task with a continuous cost for time. *Journal of Experimental Psychology*, 1971, **88**, 67–81.

Tippett, L. H. C. On the extreme individuals and the range of samples taken from a normal population. *Biometrika*, 1925, **17**, 364–387.

Vaughan, H. G., Costa, L. D., & Gilden, L. The functional relation of visual evoked response and reaction time to stimulus intensity. *Vision Research*, 1966, **6**, 645–656.

APPENDIX A

Proofs of Equations for Models with Strong Signals

Proof that p(I|s$_I$, s$_2$) *decreases with decreasing* k.

From the assumption

$$\frac{\mu_1}{\mu_2} > \frac{1 + \left(\dfrac{1}{2k-1}\right)^{\frac{1}{2}}}{1 - \left(\dfrac{1}{2k-1}\right)^{\frac{1}{2}}}$$

we conclude immediately that $\rho > \frac{1}{2}\left[1 + \left(\frac{1}{2k-1}\right)^{\frac{1}{2}}\right]$. Thus, for $x \geq \rho$,

$$\frac{\dfrac{(2k-1)!}{[(k-1)!]^2}\, x^{k-1}(1-x)^{k-1}}{\dfrac{(2k-3)!}{[(k-2)!]^2}\, x^{k-2}(1-x)^{k-2}} = 2\left(\frac{2k-1}{k-1}\right) x(1-x) < 1.$$

So

$$\begin{aligned} p_k(1|s_1,s_2) &= 1 - \frac{(2k-1)!}{[(k-1)!]^2}\int_\rho^1 [x(1-x)]^{k-1}dx > 1 - \frac{(2k-3)!}{[(k-1)!]^2}\int_\rho^1 [x(1-x)]^{k-1}dx \\ &= p_{k-1}(1|s_1,s_2). \end{aligned}$$

Proof of Equation 9

Let f_i denote the density of $\mathbf{T}_i$, then

$$P(\mathbf{T}_2 - \mathbf{T}_1 \geq \epsilon) = \int_\epsilon^\infty \int_{\max(0,-t)}^\infty f_2(x+t) f_1(x)\,dx\,dt$$

$$= \frac{(\mu_1\mu_2)^k}{[(k-1)!]^2} \int_\epsilon^\infty e^{-\mu_2 t} \int_{\max(0,-t)}^\infty [x(x+t)]^{k-1} e^{-(\mu_1+\mu_2)x}\,dx\,dt.$$

Observe that

$$\int_0^\infty [x(x+t)]^{k-1} e^{-(\mu_1+\mu_2)x}\,dx = \int_0^\infty x^{k-1} \sum_{i=0}^{k-1} \binom{k-1}{i} x^i t^{k-1-i} e^{-(\mu_1+\mu_2)x}\,dx$$

$$= \sum_{i=0}^{k-1} \binom{k-1}{i} t^{k-1-i} \frac{(k-1+i)!}{(\mu_1+\mu_2)^{k+i}}.$$

For $\epsilon \geq 0$, $t \geq 0$ and so $\max(0, -t) = 0$, and substituting and simplifying yields the first part of Equation 9. The second part, $\epsilon < 0$, requires separating the integral at 0,

$$P(\mathbf{T}_2 - \mathbf{T}_1 \geq \epsilon) = \frac{(\mu_1\mu_2)^k}{[(k-1)!]^2} \left\{ \int_\epsilon^0 e^{-\mu_2 t} \int_{-t}^\infty [x(x+t)]^{k-1} e^{-(\mu_1+\mu_2)x}\,dx\,dt \right.$$

$$\left. + \int_0^\infty e^{-\mu_2 t} \int_0^\infty [x(x+t)]^{k-1} e^{-(\mu_1+\mu_2)x}\,dx\,dt \right.$$

$$= \frac{(\mu_1\mu_2)^k}{[(k-1)!]^2} \int_\epsilon^0 e^{\mu_1 t} \sum_{i=0}^{k-1} \binom{k-1}{i} t^{k-1-i} \frac{(k+i-1)!}{(\mu_1+\mu_2)^{k+i}}\,dt + P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0)$$

$$= (1-\rho)^k \sum_{i=0}^{k-1} \binom{k+i-1}{i} \rho^k \left[1 - e^{\mu_1 \epsilon} \sum_{J=0}^{k-1-i} \frac{(-\mu_1\epsilon)^j}{j!} \right.$$

$$\left. + P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0) \right]$$

Proof of Equation 12

The proof of Equation 12 parallels that of Equation 5:

$$P\left(\frac{\mathbf{T}_2}{\mathbf{T}_1} \geq \delta\right) = \int_\delta^\infty P\left(\frac{\mathbf{T}_2}{\mathbf{T}_1} = t\right) dt$$

$$= \int_\delta^\infty \int_0^\infty f_2(ty) y f_1(y)\,dy\,dt$$

$$= \frac{(2k-1)!}{[(k-1)!]^2} \int_\delta^\infty \frac{\left(\frac{\mu_2}{\mu_1}\right)^k t^{k-1}}{\left(1 + \frac{\mu_2}{\mu_1} t\right)^{2k}}\,dt,$$

which yields Equation 12 by substituting

$$x = \frac{1}{1 + \frac{\mu_2}{\mu_1} t}.$$

APPENDIX B

Proofs of Equations for Models with Weak Signals

Proofs of Equations 33–37

Assume $J = \kappa = 1$, a signal of duration T, and the response rule that the first IAT is compared with β and respond Y if it is $\leq \beta$; N otherwise. Let $\mathbf{Y}_i$ denote the time of the ith pulse. For $\beta \leq$ T,

$$l_{s\mathrm{Y}}(t) = \frac{1}{p(\mathrm{Y}|s)} \int_{\max(0,t-\beta)}^{t} P(\mathbf{Y}_2 = t|\mathbf{Y}_1 = x\&s)\mathbf{P}(\mathrm{Y}_1 = x|s)dx$$

$$= \frac{1}{p(\mathrm{Y}|s)} \begin{cases} \int_0^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx, & t < \beta \\ \int_{t-\beta}^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx. & \beta \leq t < T \\ \int_{t-\beta}^T \mu e^{-\mu x}e^{-\mu(T-x)}\nu e^{-\nu(t-T)}dt + e^{-\mu T}\int_T^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T \leq t < T+\beta \\ e^{-\mu T}\int_{t-\beta}^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(\mathrm{Y}|s)} \begin{cases} \mu^2 te^{-\mu t}, & t < \beta \\ \mu^2\beta e^{-\mu t}, & \beta \leq t < T \\ \nu e^{-(\mu-\nu)T}e^{-\nu t}[(\mu-\nu)(T-t)+\mu\beta], & T \leq t < T+\beta \\ \nu^2\beta e^{-(\mu-\nu)T}e^{-\nu t}, & T+\beta \leq t. \end{cases}$$

Similarly,

$$l_{s\mathrm{N}}(t) = \frac{1}{p(\mathrm{Y}|s)} \begin{cases} 0, & t < \beta \\ P(\mathbf{Y}_1 = t-\beta|s)P(\mathbf{Y}_2 - \mathbf{Y}_1 \geq \beta|\mathrm{Y}_1 = t-\beta\&s) & t \geq \beta \end{cases}$$

$$= \frac{1}{p(\mathrm{N}|s)} \begin{cases} 0, & t < \beta \\ \mu e^{-\mu(t-\beta)}e^{-\mu\beta}, & \beta \leq t < T \\ \mu e^{-\mu(t-\beta)}e^{-\mu(T-t+\beta)}e^{-\nu(t-T)}, & T \leq t < T+\beta \\ e^{-\mu T}\nu e^{-\nu(t-\beta-T)}e^{-\nu\beta}, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(\mathrm{N}|s)} \begin{cases} 0, & t < \beta \\ \mu e^{-\mu t}, & \beta \leq t < T \\ \mu e^{-(\mu-\nu)T}e^{-\nu t}, & T \leq t < T+\beta \\ \nu e^{-(\mu-\nu)T}e^{-\nu t}, & T+\beta \leq t. \end{cases}$$

The argument for $\beta > T$ is similar:

$$l_{s\mathrm{Y}}(t) = \frac{1}{p(\mathrm{Y}|s)} \begin{cases} \int_0^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx, & t < T \\ \int_0^T \mu e^{-\mu x}e^{-\mu(T-x)}\nu e^{-\nu(t-T)}dx + \int_T^t e^{-\mu T}\nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T \leq t < \beta \\ \int_{t-\beta}^T \mu e^{-\mu x}e^{-\mu(T-x)}\nu e^{-\nu(t-T)}dx + e^{-\mu T}\int_T^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & \beta \leq t < T+\beta \\ e^{-\mu t}\int_{t-\beta}^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(\mathrm{Y}\,|\,s)} \begin{cases} \mu^2 t e^{-\mu t}, & t < T \\ \mu\nu e^{-(\mu-\nu)T} T e^{-\nu t} + \nu^2 e^{-(\mu-\nu)T}(t - T)e^{-\nu t}, & T \le t < \beta \\ \mu\nu e^{-(\mu-\nu)T}(T - t + \beta)e^{-\nu t} + \nu^2 e^{-(\mu-\nu)T}(t - T)e^{-\nu t}, & \beta \le t < T + \beta \\ \nu^2 \beta e^{-(\mu-\nu)T} e^{-\nu t}, & T + \beta \le t. \end{cases}$$

$$l_{s\mathrm{N}}(t) = \frac{1}{p(\mathrm{N}\,|\,s)} \begin{cases} 0, & t < \beta \\ \mu e^{-\mu(t-\beta)} e^{-\mu(T-t+\beta)} e^{-\nu(t-T)}, & \beta \le t < T + \beta \\ e^{-\mu T} \nu e^{-\nu(t-\beta-T)} e^{-\nu\beta}, & T + \beta \le t \end{cases}$$

$$= \frac{1}{p(\mathrm{N}\,|\,s)}\, e^{-(\mu-\nu)T} e^{-\nu t} \begin{cases} 0, & t < \beta \\ \mu, & \beta \le t < T + \beta \\ \nu, & T + \beta \le t. \end{cases}$$

Since these are densities, the expressions for $p(\mathrm{Y}\,|\,s)$ and $p(\mathrm{Y}\,|\,n)$ may be obtained by integrating them and setting that $= 1$. Equations 33 and 34 result. To obtain Equations 36 and 37 for MRTs, we first calculate the Fourier transforms:

$$p(\mathrm{Y}\,|\,s) L_{s\mathrm{Y}}(\omega) = \left(\frac{\mu}{i\omega + \mu}\right)^2 (1 - e^{(i\omega+\mu)\beta}) - \left(\frac{\mu}{i\omega + \mu}\right) \mu\beta e^{-(i\omega+\mu)T}$$
$$+ \left(\frac{\nu}{i\omega + \nu}\right) e^{-(i\omega+\mu)T} \left[\mu\beta - \left(\frac{\mu - \nu}{i\omega + \nu}\right)(1 - e^{-(i\omega+\nu)\beta})\right]$$

$$p(\mathrm{N}\,|\,s) L_{s\mathrm{N}}(\omega) = \left(\frac{\mu}{i\omega + \mu}\right)(e^{-(i\omega+\mu)\beta} - e^{-(i\omega+\mu)T}) + \frac{e^{-(i\omega+\mu)T}}{i\omega + \nu}[\mu - (\mu - \nu)e^{-(i\omega+\nu)\beta}].$$

We then use the fact that

$$\int_0^\infty t f(t)\,dt = i \lim_{\omega \to 0} \frac{dF(\omega)}{d\omega}$$

to obtain the MRTs.

Proofs of Equations 38–40

Observe that no response is initiated later than T and, since the residual latency is bounded by τ, no response takes longer than $\tau + T$.

The calculation of Y responses is much as before:

$$l_{s\mathrm{Y}}(t) = \frac{1}{p(\mathrm{Y}\,|\,s)} \begin{cases} \displaystyle\int_0^t \mu e^{-\mu x} \mu e^{-\mu(t-x)}\,dx & t < \beta \\ \displaystyle\int_{t-\beta}^t \mu e^{-\mu x} \mu e^{-\mu(t-x)}\,dx, & \beta \le t < T \\ 0 & T \le t \end{cases}$$

$$= \frac{1}{p(\mathrm{Y}\,|\,s)} \begin{cases} \mu^2 t e^{-\mu t}, & t < \beta \\ \mu^2 \beta e^{-\mu t}, & \beta \le t < T \\ 0, & T \le t. \end{cases}$$

The N responses arise in two ways: either the first IAT is wholly within $(0,T)$ and it exceeds β or there is at most one pulse in $(0,T)$:

$$l_{s\mathrm{N}}(t) = \frac{1}{p(\mathrm{N}\,|\,s)} \begin{cases} 0 & t < \beta \\ \mu e^{-\mu(t-\beta)} e^{-\mu\beta} = \mu e^{-\mu t} & \beta \le t < T \\ (1 + \mu\beta)e^{-\mu T} & t = T \\ 0 & T < t. \end{cases}$$

Note that the expressions for n trials are obtained by replacing μ by ν.

By integrating either density (the latter is easier), we obtain Equation 38. The Fourier transforms are:

$$p(\mathrm{Y}\,|\,s)L_{s\mathrm{Y}}(\omega) = \left(\frac{\mu}{i\omega+\mu}\right)^2 (1 - e^{-(i\omega+\mu)\beta}) - \left(\frac{\mu^2\beta}{i\omega+\mu}\right)e^{-(i\omega+\mu)T}$$

$$p(\mathrm{N}\,|\,s)L_{s\mathrm{N}}(\omega) = \left(\frac{\mu}{i\omega+\mu}\right)(e^{-(i\omega+\mu)\beta} - e^{-(i\omega+\mu)T}) + (1+\mu\beta)e^{-(i\omega+\mu)T},$$

from which Equations 39 and 40 follow.

APPENDIX C

PROOFS OF EQUATIONS FOR SIMPLE REACTION TIMES

Given the definitions of l and l_0, Equations 43 and 44, and the independent exponential distribution of IATs, it is not difficult to see that they satisfy the following pair of functional-integral equations:

$$l(x,t) = \begin{cases} \displaystyle\int_0^t \nu e^{-\nu y} l_0(x-y,\, t-y)dy, & t < x \\ \displaystyle\int_0^x \nu e^{-\nu y} l_0(x-y,\, t-y)dy + e^{-\nu x}\int_x^t \mu e^{-\mu(y-x)} l_0(0,\, t-y)dy, & t \geq x \end{cases} \tag*{[C1]}$$

$$l_0(x,t) = \begin{cases} \left.\begin{array}{ll} \nu e^{-\nu t}, & t < x \\ e^{-\nu x}\mu e^{-\mu(t-x)} & x \leq t \leq \beta \\ e^{-\nu x}e^{-\mu(\beta-x)} l(0,\, t-\beta), & x < \beta \leq t \end{array}\right\} x < \beta \\ \left.\begin{array}{ll} \nu e^{-\nu t} & t \leq \beta \\ e^{-\nu\beta} l(x-\beta,\, t-\beta), & \beta < t \end{array}\right\} x \geq \beta. \end{cases} \tag*{[C2]}$$

Setting $x = 0$ in Equation C2 yields Equation 45.
For $x \leq \beta,\, t$, Equation C2 yields

$$\begin{aligned} l_0(x,t) &= \begin{cases} e^{-\nu x}\mu e^{-\mu(t-x)}, & x \leq t \leq \beta \\ e^{-\mu x}e^{-\mu(\beta-x)} l(0,\, t-\beta), & x \leq \beta < t \end{cases} \\ &= e^{-\nu x}e^{\mu x}\begin{cases} \mu e^{-\mu t} \\ e^{-\mu\beta} l(0,\, t-\beta) \end{cases} \\ &= e^{(\mu-\nu)x} l_0(0,t). \end{aligned}$$

Using this in Equation C1 yields for $x \leq \beta$,

$$l(x,t) = \begin{cases} \nu^2 t e^{-\nu t}, & t < x \\ e^{(\mu-\nu)x}\left[\nu\displaystyle\int_0^x e^{-\mu y} l_0(0,\, t-y)dy + \mu\int_x^t e^{-\mu y} l_0(0,\, t-y)dy\right], & t \geq x. \end{cases} \tag*{[C3]}$$

Proof of Equation 46

Define

$$\varphi(t) = l_0(0,t)e^{\mu t}/\mu.$$

By Equation 45, φ satisfies

$$\varphi(t) = \begin{cases} 1, & 0 \leq t < \beta \\ \mu \displaystyle\int_0^{t-\beta} \varphi(y)dy, & \beta \leq t. \end{cases}$$

We show that

$$\varphi(t) = \sum_{i=0}^{m-1} a(m-1-i)\frac{\mu^i[t-(m-1)\beta]^i}{i!}, \quad (m-1)\beta \leq t < m\beta,$$

solves this equation if a is defined recursively as in Equation 47. Since this function is 1 for $0 \leq t < \beta$, we consider $t \geq \beta$. Suppose m is such that $(m-1)\beta \leq t < m\beta$, then by induction

$$\begin{aligned} \mu\int_0^{t-\beta} \varphi(z)dz &= \mu\left[\sum_{i=1}^{m-2}\int_{(i-1)\beta}^{i\beta} \varphi(z)dz + \int_{(m-2)\beta}^{t-\beta} \varphi(z)dz\right] \\ &= \mu\left\{\sum_{i=1}^{m-2}\int_0^{\beta} \varphi[y+(i-1)\beta]dy + \int_0^{t-(m-1)\beta} \varphi[y+(m-2)\beta]dy\right\} \\ &= \mu\left\{\sum_{i=1}^{m-2}\int_0^{\beta}\sum_{j=0}^{L-1} a(i-1-j)\frac{\mu^j y^j}{j!}dy + \int_0^{t-(m-1)\beta}\sum_{j=0}^{m-2} a(m-2-j)\frac{\mu^j y^j}{j!}dy\right\} \\ &= \sum_{i=1}^{m-2}\sum_{k=1}^{i} a(i-k)\frac{(\mu\beta)^k}{k!} + \sum_{k=1}^{m-1} a(m-1-k)\frac{\mu^k[t-(m-1)\beta]^k}{k!} \\ &= a(m-1) + \sum_{k=1}^{m-1} a(m-1-k)\frac{\mu^k[t-(m-1)\beta]^k}{k!} \\ &= \varphi(t). \end{aligned}$$

Another approach to solving integral equations of retarded type is to differentiate them and attempt to reduce the problem to one of differential equations of retarded type. These have been studied in some detail (see Bellman & Cooke, 1963; El'sgol'ts, 1966).

Care must be taken, however, since information is lost. For example, our integral equation for φ says φ is discontinuous at $t = \beta$. If we differentiate, ignoring all potential discontinuities,

$$\begin{aligned} \varphi'(t) &= \mu\varphi(t-\beta), & t > \beta, \\ \varphi(t) &= 1, & 0 \leq t < \beta. \end{aligned}$$

This is dealt with on page 6 on El'sgol'ts and the solution is

$$\varphi(t) = \sum_{i=0}^{m-1} \mu^i(t-i\beta)^i/i!, \quad (m-1)\beta \leq t < m\beta.$$

It is easy to see that this does not solve the integral equation because it is continuous at $t = \beta$.

Proof of Equation 48

Rewrite Equation 45 and use the linear term of the Taylor series,

$$\begin{aligned} e^{\mu t} l_0(0,t) &= \mu\begin{cases} 1, & t \leq \beta \\ \displaystyle\int_0^{t-\beta} e^{\mu y} l_0(0,y)dy & t > \beta \end{cases} \\ &\cong \mu\begin{cases} 1 \\ \displaystyle\int_0^{t} e^{\mu y} l_0(0,y)dy - \beta e^{\mu t} l_0(0,t). \end{cases} \end{aligned}$$

Differentiate this for $t > \beta$ and collect terms

$$l'_0(0,t) = -\mu' l_0(0,t),$$

where

$$\mu' = \frac{\mu^2\beta}{1+\mu\beta}.$$

Equation 48 follows by integration except for a constant of integration which is evaluated by setting $\int_0^\infty l_0(0,t)dt = 1$.

Proof of Equation 50

From Equation 45,

$$\begin{aligned} L_0(0,\omega) &= \int_0^\infty e^{-i\omega t} l_0(0,t)dt \\ &= \int_0^\beta e^{-i\omega t}\mu e^{-\mu t}dt + \int_\beta^\infty e^{-i\omega t}\mu e^{-\mu t}\int_0^{t-\beta} e^{\mu y} l_0(0,y)dy \\ &= \frac{\mu}{i\omega+\mu}[1 - e^{-(i\omega+\mu)\beta}] + \mu e^{(i\omega+\mu)\beta}\int_0^\infty e^{-(i\omega+\mu)z}\int_0^z e^{\mu y} l_0(0,y)dy. \end{aligned}$$

Integrating by parts, taking into account that $l_0(0,t) \to 0$ as $t \to \infty$ because l_0 is a density, we obtain

$$L_0(0,\omega) = \frac{\mu}{i\omega+\mu}[1 - e^{-(i\omega+\mu)\beta} + e^{-(i\omega+\mu)\beta}L_0(0,\omega)].$$

Solving yields Equation 50.

Proof of Equation 51

By Equation C3,

$$\begin{aligned} L(x,\omega) &= \int_0^x e^{-i\omega t} l(x,t)dt + \int_x^\infty e^{-i\omega t} l(x,t)dt \\ &= \int_0^x e^{-i\omega t}\nu^2 t e^{-\nu t}dt + e^{(\mu-\nu)x}\int_x^\infty e^{-i\omega t}\nu\int_0^x e^{-\mu y} l_0(0, t-y)dydt \\ &\qquad + e^{(\mu-\nu)x}\int_x^\infty e^{-i\omega t}\mu\int_x^t e^{-\mu y} l_0(0, t-y)dydt. \end{aligned}$$

The first integral is trivial. The second one can be dealt with by a change of variable and an integration by parts:

$$\begin{aligned} \nu e^{(\mu-\nu)x}\int_x^\infty e^{-(i\omega t+\mu)t}\int_{t-x}^t e^{\mu z} l_0(0,z)dzdt &= \frac{\nu e^{(\mu-\nu)x}}{i\omega+\mu}\Bigg\{ - e^{-(i\omega+\mu)t}\int_{t-x}^t e^{\mu z} l_0(0,z)dz\Bigg|_x^\infty \\ &\qquad + \int_x^\infty e^{-(i\omega+\mu)t}[e^{\mu t} l_0(0,t) - e^{\mu(t-x)} l_0(0, t-x)]dt\Bigg\} \\ &= \frac{\nu e^{(\mu-\nu)x}}{i\omega+\mu}\Bigg\{\mu x e^{-(i\omega+\mu)x} \\ &\qquad + \left[L_0(0,\omega) - \frac{\mu}{i\omega+\mu}\right][1 - e^{-(i\omega+\mu)x}]\Bigg\}. \end{aligned}$$

A similar treatment of the third integral yields

$$\frac{\mu e^{(\mu-\nu)x}e^{-(i\omega+\mu)x}}{i\omega+\mu}L_0(0,\omega).$$

Adding the three terms gives Equation 51.

Proof of Equation 54

In the following estimates of $l(x,t)$ we are only concerned with $t \geqslant x$ and so it will prove convenient to write the second argument as $t + x$, $t \geqslant 0$. Consider, first, the case $x \leqslant \beta$. The plan is to substitute the approximate solution for $l_0(0,t)$, Equation 48, into the second part of Equation C3. Observe that if we confine ourselves to $t > \beta$, then $t + x - y \leqslant \beta$ iff $y \geqslant t + x - \beta > x$, hence the two integrals of Equation C3 divide into three:

$$l(x, t+x) \cong e^{(\mu-\nu)x}\left\{\nu\int_0^x e^{-\mu y}\mu' e^{-\mu'/\mu}e^{-\mu'(t+x-y)}dy + \mu\int_x^{t+x-\beta} e^{-\mu y}\mu' e^{-\mu'/\mu}e^{-\mu'(t+x-y)}dy \right.$$
$$\left. + \mu\int_{t+x-\beta}^{t+x} e^{-\mu y}\mu e^{-\mu(t+x-y)}dy\right\}.$$

Carrying out the integrations and taking note of the fact that $\mu = \mu' + \mu/(1 + \mu\beta)$ yields

$$l(x, x+t) \cong A''e^{-\mu t} + B''e^{-\mu' t}, \quad x \leqslant \beta < t,$$

where A'', B'' are functions of ν, μ, β, and x which we do not need to specify.

For $x > \beta$, we proceed by a somewhat different argument.[12] Let $\mathbf{Y}$ be the time of the last pulse, if any, prior to the signal onset at time x, or 0 otherwise, and let $\mathbf{Y}'$ be the time of the first pulse after the signal onset. Since, in a Poisson process, the time from x back to the preceding pulse is exponentially distributed with the parameter of the process, we see that

$$P(\mathbf{Y}' = y, \mathbf{Y}' - \mathbf{Y} > \beta \,|\, \mathbf{S} = x > \beta) = \mu e^{-\mu y}\begin{cases} e^{-\nu(\beta-y)}, & y \leqslant \beta \\ 1, & y > \beta. \end{cases}$$

Observe that if $t > 2\beta$, then $t - y \leqslant \beta$ iff $y \geqslant t - \beta > \beta$, so

$$\begin{aligned} l(x, t+x) &= \int_0^t P(\mathbf{Y}' = y, \mathbf{Y}' - \mathbf{Y} > \beta \,|\, \mathbf{S} = x > \beta)\, l_0(0, t-y)dy \\ &\cong \int_0^\beta \mu e^{-\mu y}e^{-\nu(\beta-y)}\mu' e^{-\mu'/\mu}e^{-\mu'(t-y)}dy + \int_\beta^{t-\beta} \mu e^{-\mu y}\mu' e^{-\mu'/\mu}e^{-\mu'(t-y)}dy \\ &\qquad + \int_{t-\beta}^{t} \mu e^{-\mu y}\mu e^{-\mu(t-y)}dy \\ &= A''e^{-\mu t} + B''e^{-\mu' t}, \quad x > \beta, \quad t > 2\beta, \end{aligned}$$

where A'', B'' are functions of ν, μ, and β in the range $x > \beta$.

Thus, the density of RTs $\geqslant \tau$ to a signal with onset x is

$$f(x, t+x) = \int_0^\tau l(x, t+x-y)r(y)dy, \quad t \geqslant \tau.$$

For $t > \tau + 2\beta$, $t - y > 2\beta$ and so we may substitute the above approximations for $l(x, t + x - y)$, yielding

$$f(x, t+x) = A'e^{-\mu t} + B'e^{-\mu' t}, \quad \text{for} \quad t > \tau + 2\beta,$$

---

[12] We are indebted to E. Thomas for pointing out this argument.

where A', B' are functions of ν, μ, β, x, and r. Assuming an arbitrary density g of foreperiods,

$$\begin{aligned} f_{R-S}(t) &= C\int_0^\infty g(x)f(x,\, t+x)dx \\ &= Ae^{-\mu t} + Be^{-\mu' t}, \quad \text{for} \quad t > \tau + 2\beta, \end{aligned}$$

where A, B are functions of ν, μ, β, g, and r.

Proof of Equation 55

First, we show that for $t \le x$, $l(x,t) = l(t,t)$. By Equation C1, we see that it is sufficient to show that $l_0(x,t) = l_0(t,t)$. By Equation C3, this is true for $x \le \beta$ and for $t \le \beta < x$. For $\beta < t \le x$, it is true provided that $l(x-\beta,\, t-\beta) = l(t-\beta,\, t-\beta)$. By a finite induction of the form $x - i\beta$, the proof is reduced to the already proved case of $x \le \beta$. Now, with $t \le x$, Equations C1 and C2 yield

$$\begin{aligned} l(t,t) &= \int_0^t \nu e^{-\nu y} l_0(t-y,\, t-y)dy \\ &= \begin{cases} \displaystyle\int_0^t \nu e^{-\nu y}\nu e^{-(t-y)}dy, & t \le \beta \\ \displaystyle\int_0^\beta \nu e^{-\nu y}\nu e^{-\nu(t-y)}dy + \int_\beta^t \nu e^{-\nu y}e^{-\nu\beta} l(t-y-\beta,\, t-y-\beta)dy, & t > \beta \end{cases} \\ &= \begin{cases} \nu^2 t e^{-\nu t}, & t \le \beta \\ \displaystyle\nu^2\beta e^{-\nu t} + \nu e^{-\nu t}\int_0^{t-\beta} e^{\nu z} l(z,z)dz, & t > \beta. \end{cases} \end{aligned}$$

Proof of Equation 59

Assume $t \ge \tau' + \tau$. Then

$$\begin{aligned} f_R(t) &= K'\int_t^\infty \lambda e^{-\lambda x}\int_0^t l(x,\, t-y)r(y)dydx \\ &= K'\int_t^\infty \lambda e^{-\lambda x}dx\int_0^\tau l(t-y,\, t-y)r(y)dy \\ &= K'e^{-\lambda t}\int_0^\tau De^{-\nu'(t-y)}r(y)dy \\ &= K'De^{-(\lambda+\nu')t}\int_0^\tau e^{\nu' y}r(y)dy. \end{aligned}$$

Partial Results about l(x,t) *for* x $\geqslant \beta$

From the argument given in deriving Equation 54, for $x > \beta$,

$$l(x,\, t+x) = \begin{cases} \displaystyle\mu e^{-\nu\beta}e^{-(\mu-\nu)t}\int_0^t e^{(\mu-\nu)y} l_0(0,y)dy, & 0 \leqslant t \leqslant \beta \\ \displaystyle\mu e^{-\nu\beta}e^{-(\mu-\nu)t}\int_{t-\beta}^t e^{(\mu-\nu)y} l_0(0,y)dy + \mu e^{-\mu t}\int_0^{t-\beta} e^{\mu y} l_0(0,y)dy, & t > \beta. \end{cases}$$

By routine integration,

$$\int_0^\infty e^{-i\omega t} l(x,\, t + x)dt = \frac{\mu e^{-\nu\beta} L_0(0,\omega)}{\mu + i\omega},$$

where $L_0(0,\omega)$ is given by Equation 50. So, using Equation 55,

$$\begin{aligned} L(x,\omega) &= \int_0^\infty e^{-i\omega t} l(x,t)dt \\ &= \int_0^x e^{-i\omega t} l(t,t)dt + e^{-i\omega x} \int_0^\infty e^{-i\omega y} l(x,\, y + x)dy \\ &= \left(\frac{\nu}{\nu + i\omega}\right)^2 (1 - [(\nu + i\omega)\beta + 1]e^{-(\nu + i\omega)\beta}) + \int_\beta^x e^{-i\omega t} l(t,t)dt + \frac{e^{-i\omega x}\mu e^{-\nu\beta} L_0(0,\omega)}{\mu + i\omega}. \end{aligned}$$

The problem in going further is how to evaluate the integral in the second term. One possibility is to use the approximation of Equation 58; however, we are uncertain how good it is near β. If so, then it is easy to calculate the mean decision time for $x > \beta$, which together with Equation 52 allows us to calculate the MRT for any distribution of foreperiods. As this expression is both complex and approximate, we do not present it.

(Received March 30, 1971)

Psychological Review
1972, Vol. 79, No. 1, 58–67

COGNITIVE CONTROL[1]

KENNETH R. HAMMOND[2]
University of Colorado

AND

DAVID A. SUMMERS
University of Kansas

The theory developed here proposes that performance in cognitive tasks involves two distinct processes: acquisition of knowledge and cognitive control over knowledge already acquired. A conceptual and analytic framework is presented which allows for the disentanglement of knowledge and control, and for the quantification of each. Evidence from studies of multiple-cue probability learning, clinical judgment, and interpersonal conflict supports the theoretical usefulness of this framework and indicates that poor performance in cognitive tasks can often be attributed to incomplete cognitive control, rather than incomplete knowledge. The importance of cognitive feedback—as opposed to traditional outcome feedback—for the development of cognitive control is illustrated.

Although learning theorists have long emphasized the distinction between learning and performance, little attention has been given to skill in the application of knowledge in tasks which do not involve motor performance. Rather, there is an implicit assumption that once knowledge has been acquired, the application of this knowledge is largely dependent on certain experimental circumstances.

For example, Deese and Hulse (1967) point out that

> sometimes we have good reason to believe that an organism is simply not demonstrating what it has learned because we have not chosen the proper conditions which will assure the overt display of the appropriate behavior [p. 62].

Leeper (1970) makes a similar point, but draws attention specifically to motivational factors:

> One of the distinctive points of cognitive learning theory—a principle developed by Lashley (1929), Tolman (1932), and others—is that a clear distinction must be made between learning (or the "acquisition of habits") and performance (or the "use of habits"), and that a good share of the "use of habits" depends on current motivation [p. 287].

Other theorists, also stressing the importance of distinguishing between learning and performance, have suggested additional determinants of the application of knowledge; for example, fatigue (Hilgard, 1956) and instructions (Mandler, 1967). Indeed, Mandler goes somewhat further and suggests that a distinction be made between "the conditions under which a subject will discover a particular rule [and] his *abilities* to use . . . it [p. 22; italics added]."

Although the explanations for the inappropriate application of knowledge mentioned above are by no means identical, they are (with the exception of Mandler) quite similar in one important respect. Specifically, they all assume that the potential for full *application* of knowledge accompanies full *acquisition* of knowledge—except in psychomotor tasks.

The position taken here, however, is that acquisition and application are independent components of learning in cognitive tasks as well as psychomotor tasks. Consequently, we argue that the development of skill—or control—in the application of knowledge already acquired is a matter to be investigated. The purpose of this paper, therefore, is to introduce the concept of cognitive control, to indicate its theoretical and methodological context, and to illustrate its empirical significance in studies of human learning, judgment, and interpersonal behavior. In addition, traditional outcome feedback is identified as an impediment to the development of cognitive control.

---

[1] This paper is based on research undertaken at the Institute of Behavioral Science, University of Colorado, and is Publication 140 of the Institute. The research was supported by National Institute of Mental Health Grant MH-16437-02. The authors wish to thank Thomas Stewart for his assistance. We are particularly indebted to Berndt Brehmer of the University of Umeå for his valuable contributions made while he was a visiting Research Associate at the University of Colorado.

[2] Requests for reprints should be sent to Kenneth R. Hammond, Institute of Behavioral Science, University of Colorado, Boulder, Colorado 80302.

THEORY

Fundamental to our treatment of cognitive control is a methodological and conceptual framework developed for studying behavior in multiple-cue probability learning tasks. A major reason for the emphasis on multiple-cue probability learning tasks is the fact that such tasks have essential features in common with a broad range of inference tasks human subjects routinely encountered outside the laboratory. Aside from this, there are two additional reasons for studying the problem of cognitive control in the framework of multiple-cue probability learning tasks: (*a*) Multiple-cue probability learning tasks can be varied sufficiently in complexity and uncertainty so that both knowledge and cognitive control can be varied extensively, and (*b*) quantitative separation of knowledge and control has been accomplished within this framework.

Variations in Complexity and Uncertainty

Multiple-cue probability learning tasks can be varied in both complexity and uncertainty in at least three ways: (*a*) the number of cues related to a criterion can be varied, (*b*) the uncertainty associated with each cue can be varied by creating differential cue validities, and (*c*) the form of the relation between cue and criterion can be varied; either positive linear relations, or various non-negative linear relations, or various nonlinear relations may be used (see Brehmer & Lindberg, 1970; Knowles, Hammond, Stewart, & Summers, 1971; Naylor & Clark, 1968; Sheets & Miller, in press; Summers, 1969; Summers & Hammond, 1966; Uhl, 1963).

Quantitative Denotation of Knowledge and Control

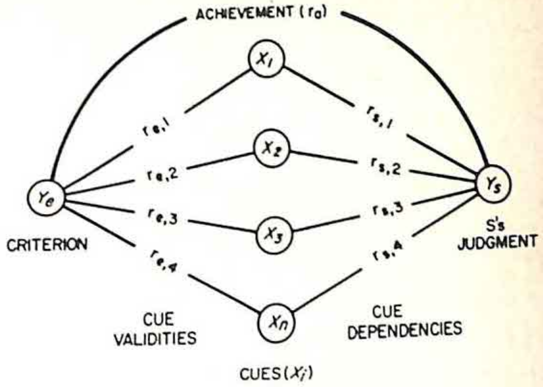


FIG. 1. Brunswik's lens model.

The analysis of multiple-cue probability learning has typically been carried out within the Brunswikian framework (Brunswik, 1956) as modified by Hammond (1966). As illustrated in Figure 1, Brunswik's "lens model" has been used to convey a pictorial model of the relation between subject and task. The mathematical specification of the relation between cue validities and cue dependencies in Figure 1 was introduced by Hursch, Hammond, and Hursch (1964), and modified in various ways subsequently (Dudycha & Naylor, 1966; Rozeboom, 1971; Tucker, 1964). Tucker's (1964) development of the lens model equation will be used here. It reads as follows:

$$r_a = GR_eR_s + C\sqrt{1-R_e^2}\sqrt{1-R_s^2} \qquad [1]$$

where

r_a = the correlation between Y_e and Y_s;
G = the correlation between the linear prediction of Y_e and Y_s from the cue values;
R_e = the multiple correlation between the cues and Y_e;
R_s = the multiple correlation between the cues and Y_s;
C = the correlation between the variance in the task system and the subject's judgmental system which is unaccounted for by the linear component G.

When the systematic variance in the criterion can be accounted for by a linear function of the cue values, then the contribution of the second term on the right-hand side of Equation 1 will be negligible. Equation 1 then reduces to

$$r_a = GR_eR_s \qquad [2]$$

Equation 2 may, of course, be applied to nonlinear functions in the case where the appropriate transformation can be effected—as in all the studies described below. It is Equation 2 that makes the distinction between knowledge and control most apparent. Therefore, each term in Equation 2 is discussed separately below.

r_a (achievement): This term measures performance; specifically, r_a represents the covariation of the subject's judgments (Y_s) with the criterion (Y_e).

G (knowledge): This term measures the extent to which the subject has correctly detected properties of the task. Specifically, this term reflects the covariation between the least-squares prediction of the criterion ($\hat{Y}_e$) and the least-squares prediction of the subject's judgments ($\hat{Y}_s$) from the cues. When systematic task variance can be identified and represented in a regression equation, G denotes the degree to which the subject's cognitive system is isomorphic with the task system independent of uncertainty in the task system (R_e) or control in the subject's response system (R_s).

Isomorphism as used here does not, of course, imply identity; G may reach unity, as it should, when the regression coefficients in the task system and the subjects cognitive system are proportional (see Castellan, 1971).

R_e (task uncertainty): This term measures the predictability of the criterion (Y_e) from the cues (X_1) in the task system. It sets a limit on the extent to which achievement (r_a) may occur, even if knowledge (G) and control (R_s) are perfect (see Hursch et al., 1964, also Hammond, Hursch, & Todd, 1964, for a discussion of the limits of achievement under conditions of uncertainty). R_e, of course, may be determined by the experimenter.

R_s (cognitive control): This term measures the extent to which the subject controls the execution of his knowledge; it indicates the predictability of the subject's response (Y_s) from the cues (X_1). Note that R_s is statistically independent of G. Such independence is critical, for it means that even should G reach unity (indicating perfect knowledge), if R_s were less than unity (indicating imperfect control), performance would be less than the limit of achievement (R_e) would permit. Conversely, R_s might equal 1.00, thus indicating that the perfectly controlled cognitive system was not appropriate to the task system, thus preventing achievement (r_a) from reaching its upper limit (R_e).

Two subjects, therefore, might have identical achievement indexes for different reasons; one because of perfect knowledge ($G = 1.00$) but imperfect control ($R_s < 1.00$) and the other because of perfect control ($R_s = 1.00$), but imperfect knowledge ($G < 1.00$). Variations between these extremes could also occur, of course.

APPLICATIONS OF THE THEORY

Applications of the theory will be illustrated in regard to (*a*) individual learning in complex, multiple-cue inference tasks, (*b*) clinical judgment, and (*c*) conflict between two persons working together in a multiple-cue probability learning task.

Individual Multiple-Cue Probability Learning

In a recent experiment (Deane, Hammond, & Summers, 1972), subjects were studied in two three-cue probabilistic tasks: one in which the relation between each cue and the criterion was *linear*, and one in which the relation between each cue and the criterion was *nonlinear*. It should be noted that previous research involving nonlinear inference tasks has shown the learning of nonlinear relations to occur slowly, if at all (see Brehmer, 1969; Hammond & Summers, 1965). Typically, the poor performance observed in complex nonlinear tasks is attributed to an inability of most subjects to detect the task relation—in short, poor performance is usually explained in terms of the subject's inability to acquire full knowledge of these tasks.

In the effort to evaluate the role of control, however, Deane et al. (1972) provided subjects in both the linear and nonlinear task conditions (after 20 warm-up trials) with the knowledge necessary for highly

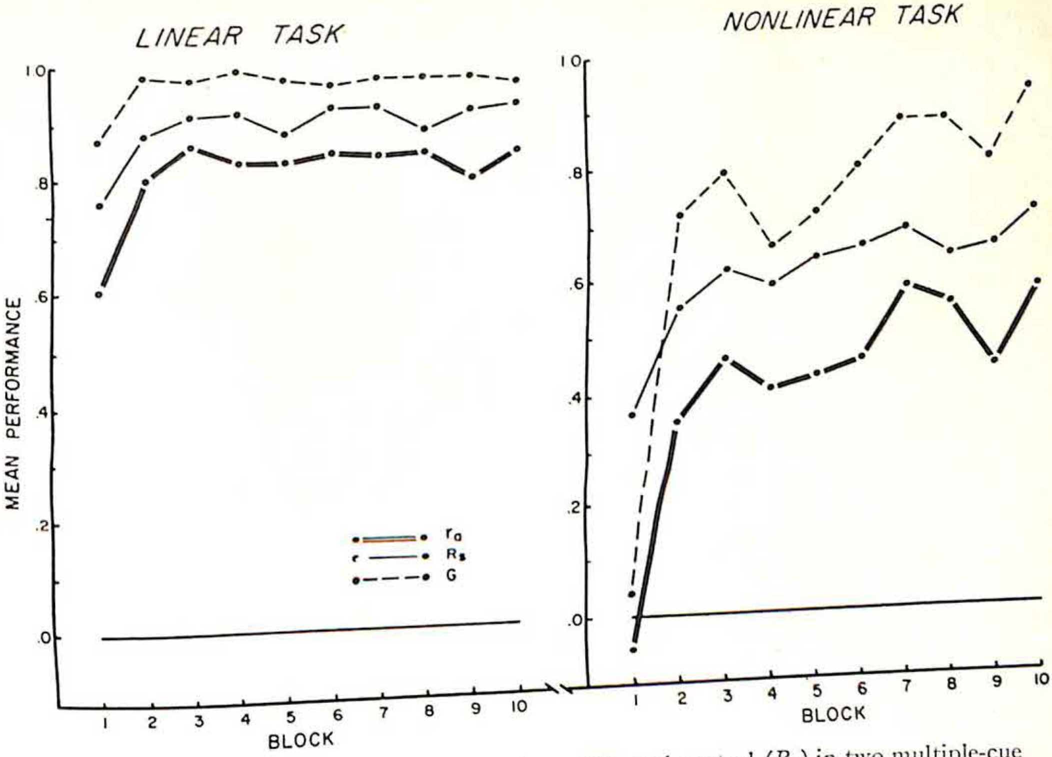


FIG. 2. Indexes of achievement (r_a), knowledge (G), and control (R_s) in two multiple-cue probability learning tasks. (In each task condition, $n = 20$.)

accurate performance; that is, subjects were told how each cue was related to the criterion dimension and how each cue should be weighted. Moreover, this information was provided to the subjects in two different ways; verbally or pictorially. Following each learning trial, the subject was given the usual outcome feedback; that is, the subject was informed of the correct answer.

The results of this experiment indicated that poor performance in complex inference tasks (e.g., those involving nonlinear relations) can be attributed to difficulties in cognitive control, as well as to difficulties in acquiring knowledge about the task. At the conclusion of training (200 trials), the level of predictive accuracy (r_a) achieved by subjects in the nonlinear task condition was significantly lower than that achieved by subjects in the linear task condition (.58 and .84, respectively). Yet, as can be seen in Figure 2, knowledge, as measured by G, was essentially the same in these two tasks by the end of training. In contrast, however, cognitive control (R_s) was far below optimal for subjects in the nonlinear task, even after 200 trials. For subjects in the linear task condition, control (R_s) was hardly distinguishable from knowledge (G).

These findings indicate that knowledge and control can be disentangled empirically, as well as statistically, and demonstrate that even when knowledge is complete, imperfect cognitive control can prevent high achievement.

It is important to note that the findings summarized above are not peculiar to tasks involving probabilistic cue-criterion relations; that is, irreducible error. For example, Brehmer (1969) studied performance in four different multiple-cue tasks—all permitting perfect accuracy ($R_e = 1.00$)—and found considerable differences according to the type of task relation the subjects were required to learn and use. When the criterion was a simple linear function of the cues ($Y_e = X_1 + X_2$), the subjects had little difficulty in rapidly achieving a high level of performance. When the task required the subjects to utilize a complex

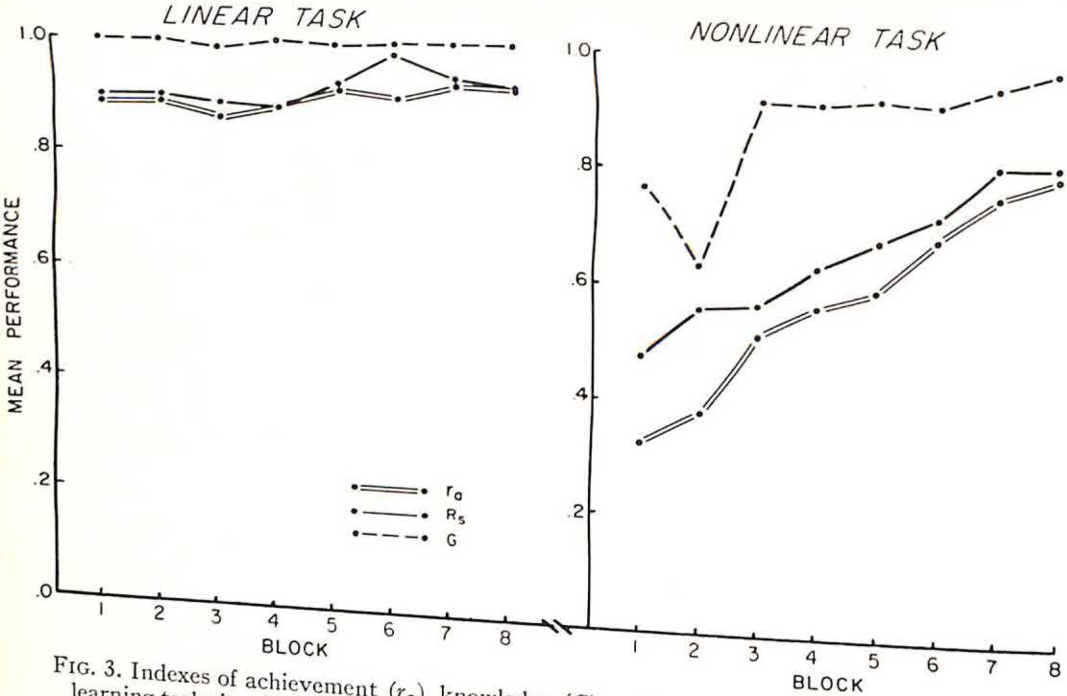


FIG. 3. Indexes of achievement (r_a), knowledge (G), and control (R_s) in two multiple-cue learning tasks in which perfect achievement is possible. (In each task condition, $n = 10$.)

nonlinear function ($Y_e = X_1/X_2$), however, performance was generally poor and improved quite slowly.

One plausible explanation for Brehmer's findings is, of course, that the subjects in the latter condition were unable to detect the task relations; that is, that subjects were unable to acquire the appropriate task knowledge. As can be seen in Figure 3, however, a reanalysis of his data using the framework proposed here suggests quite a different explanation. Note that in the linear task, both knowledge (G) and control (R_s) were high; as a consequence, performance (r_a) was near perfect. In contrast, performance (r_a) in the nonlinear task improved slowly, and remained significantly ($p < .01$) below that attained in the linear task—even at the end of 400 trials. The R_s and G curves, however, indicate that this poor performance in the nonlinear task cannot be attributed solely to the subjects inability to learn the task relations. Indeed, G improved rapidly, and at the conclusion of training was virtually indistinguishable from that attained by subjects in the linear task (.98 and .99, respectively). The primary source of the suboptimal performance in the nonlinear task was lack of full cognitive control over knowledge acquired; that is, R_s improved far more slowly than did G, and remained significantly ($p < .05$) less than that attained in the linear task at the conclusion of training (.78 and .90, respectively).

Cognitive Control in Clinical Judgment

The analyses described above illustrate the usefulness of *cognitive control* as an explanatory concept in laboratory studies of human learning. There is growing evidence that this concept is also useful in accounting for performance in nonlaboratory situations as well—such as those involving clinical judgment.

Several investigators of judgment processes (e.g., Dudycha & Naylor, 1966; Ward & Davis, 1965) have argued that judgment policies (in a variety of domains) are frequently "correct" (implying high G), but are executed in an inconsistent manner (implying low R_s). This point has been elaborated by Goldberg (1970) who made use of Tucker's version of the Lens Model Equation to demonstrate how clinical judgment is influenced by inconsistency; that is, by failure of control. Specifically, Goldberg reanalyzed the judgments of 29 clinical psychologists about 861 Minnesota Multiphasic Personality Inventory profiles and found that while G was typically high ($G = .68$), judgmental accuracy was low ($r_a = .28$).

Although there is evidence that this poor performance was partially due to an inability to detect nonlinearity in the judgment tasks, there is also evidence that it resulted from lack of control; that is, low R_s. For when Goldberg produced *perfect control* by replacing the clinical judge with a perfectly consistent linear model (in effect setting $R_s = 1.0$), performance (r_a) improved for 86% of the judges. Similar findings have been reported by Dawes (1971) in a study involving judgments made by graduate admissions committees.

These results support Dudycha and Naylor's (1966) contention that

> humans tend to generate "correct" strategies but then, in turn, fail to use their own strategy with any great consistency . . . One is left with the conclusion that humans may be used to generate inference strategies but that once the strategy is obtained the human should be removed from the system and replaced by his own strategy [p. 127].

The results also confirm our views concerning the theoretical distinction between knowledge and control.

Cognitive Control and Interpersonal Conflict

There is evidence that the framework proposed here is significant not only for individual performance, but for certain kinds of two-person interactions as well. Of particular concern here are those interpersonal conflicts (quarrels) which arise when two persons must exercise their judgments under conditions of uncertainty. Such quarrels are studied in settings in which two or more individuals who have learned to utilize probabilistic cues in quite different ways (different cue weights and/or different function forms) are required to reach joint agreement in multiple-cue probability learning tasks.

A large number of studies have investigated the conflict, or disagreement, which arise under these conditions, and all have found that the reduction of conflict occurs slowly, if at all, even when subjects are given outcome feedback which clearly indicates to each that his cognitive system is inadequate (Hammond & Brehmer, in press). A commonsense explanation suggests that the slow reduction in conflict merely reflects the reluctance of most people to alter their beliefs (even in the face of contradictory evidence). Examination of conflict behavior using the theoretical analysis proposed here, however, points to a different interpretation.

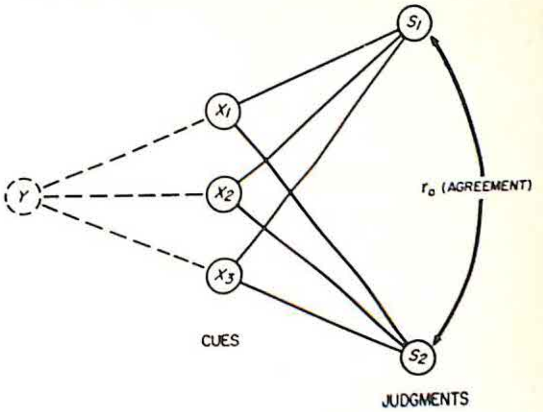


FIG. 4. Brunswik's lens model as applied to the two-person case. (Note that characteristics of the judgment task are represented on the left, while characteristics of each person's cognitive system are represented on the right.)

Such analysis is enhanced by the fact that not only can each subject's behavior be examined in relation to an environmental task system, but each subject's behavior can also be examined in relation to the response system of the other (Hammond, Wilkins, & Todd, 1966). As can be seen in Figure 4, the theoretical framework outlined earlier—including the concept of cognitive control—is directly applicable to this type of interpersonal situation as well as to the individual learning and judgment tasks described above: r_a represents agreement between the responses of the two subjects, G represents the convergence between the two systems, and R_1 and R_2 represent the cognitive control exerted by S_1 and S_2, respectively.

The results from interpersonal conflict studies conducted in eight different nations (see Brehmer, Azuma, Hammond, Kostron, & Varonos, 1970; Hammond & Brehmer, in press) have been analyzed in the above terms. The results of these studies all point in the same direction: the subjects' cognitive systems converged (*G increased*), but cognitive control diminished (*R_s decreased*) as a result of the interaction. Thus the quarrels—differences in judgments—continued despite increased agreement *in principle* (increased G); loss of cognitive control (decreased R_s) produced disagreement *in fact*. These results imply that cognitive control plays a significant part in interpersonal learning and conflict as well as in individual learning and clinical judgment.

FEEDBACK AND CONTROL

Why does cognitive control (R_s) exhibit such a slow growth curve relative to knowledge (G)? If subjects *know* what to do, why do they not apply their knowledge more effectively? The answer appears to lie in the heretofore unexamined traditional form of feedback in human learning experiments.

Outcome Feedback as an Impediment to Control.

As in the above experiments, it is customary to expect subjects to learn by providing them with the correct answer on every trial. But that tradition follows from an unexamined assumption which cognitive theory has inherited from trial and error and S-R theories of learning. Indeed, outcome feedback is still an integral part of the learning research paradigm used by all psychologists—even cognitive theorists.

Recently, however, it has become apparent that outcome feedback does not provide much help for subjects attempting to acquire knowledge about the properties of complex inference tasks, particularly when the task requires learning of complex relations under conditions of uncertainty (e.g., Todd & Hammond, 1965). Dramatic confirmation of this conclusion is provided by Goldberg (1968) who reported that even after 4,000 trials—each followed by outcome feedback—learning in a complex clinical inference task was practically nonexistent.

Of equal importance, however, are numerous findings that the removal of outcome feedback typically results in an increase in response consistency (Azuma & Cronbach, 1966; Bjorkman, 1965; Brehmer & Lindberg, 1970). These findings suggest that in addition to its limited usefulness in the acquisition of knowledge, outcome feedback may be detrimental to control, as well.

Cognitive Material as Feedback

An alternative approach, more in keeping with cognitive theory, is to provide feedback that contributes to the exercise of control rather than feedback that prevents, or at least slows, the development of control. Feedback which contributes to control should consist of *cognitive* material—not *response*-oriented material—which will enable the subjects to perceive not only that their judgment was in error, but *why* it was in error. Such cognitively oriented feedback must enable the subject to compare (*a*) the properties of his cognitive system with (*b*) the properties of the task system with which he is trying to cope, whether a content-neutral learning task or the cognitive system of another person. Of particular importance in this respect are two essential parameters: (*a*) differential cue weights and (*b*) the form of the function relating each cue to the criterion.

The critical question is: Once the subjects have been given such feedback, can they exercise sufficient cognitive control (R_s) to increase their overall performance (r_a), without loss in their grasp of the properties of the task (G)? The results of two experiments indicate that they can.

The first of these studies (Todd & Hammond, 1965) employed two different probabilistic tasks in which the subject was instructed to use the information afforded by three geometric cues to predict a numerical criterion. One of these tasks, which required the subjects to weight all three cues equally, was particularly difficult to learn; performance in this task is therefore of primary interest here.

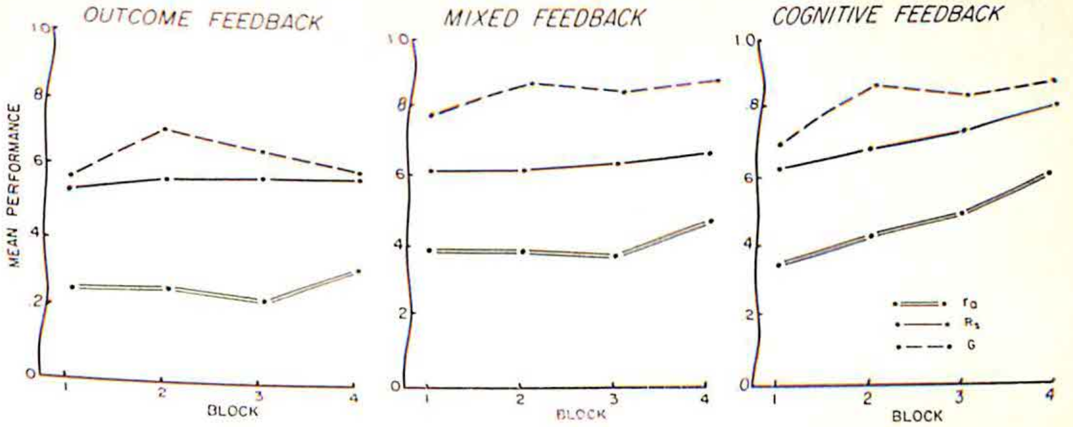


FIG. 5. Indexes of achievement (r_a), knowledge (G), and control (R_s) in a multiple-cue probability learning task under conditions of outcome, mixed, and cognitive feedback. (In each feedback condition, $n = 12$.)

The major innovation in this study is that one group of subjects received only cognitively oriented feedback during learning (referred to as "lens model" feedback by Todd and Hammond). Using an on-line computer arrangement, each subject was periodically given information about task properties (cue validities), as well as properties of his cognitive system (cue utilization coefficients). A second group was given only traditional outcome feedback; that is, the correct answer after every trial. A third group was given both cognitive and outcome feedback (referred to as "mixed" feedback).

Using r_a as a measure of performance, Todd and Hammond found that the group which received only cognitive feedback performed most accurately, followed by the mixed feedback group, and finally by the outcome feedback group. That the cognitive feedback group performed better than the outcome feedback group is theoretically gratifying; but what is surprising is that subjects who received both the cognitive and outcome feedback ("mixed" feedback) performed *less* well than subjects who were given lens model feedback only.

A reanalysis of the Todd and Hammond data revealed that performance in the mixed feedback conditions can be explicated in terms of cognitive control. As can be seen in Figure 5, subjects in both the cognitive and mixed feedback conditions were able to achieve a high degree of knowledge about the task. The G indexes for these two groups were quite high throughout, and were near identical by the end of 200 trials. On the other hand, *control* in the cognitive feedback condition was significantly ($p < .01$) higher than in the mixed feedback condition, even at the end of training. (Note that neither knowledge nor control improved substantially in the outcome feedback condition.)

In short, the cognitively oriented feedback given to subjects in the cognitive and mixed feedback conditions facilitated the acquisition of knowledge about task properties. The addition of outcome feedback to the mixed feedback condition, however, served to *decrease* cognitive control and thus contributed to the suboptimal performance achieved by these subjects.

A second study employed the same nonlinear judgment task used by Deane et al. (described earlier), but used an on-line computer graphics terminal as a means of providing feedback. This technique enabled the subject to compare periodically his cue weights and function forms with the properties of the task system in pictorial form. (Some practical implications of the rapid learning associated with computer graphics techniques are pointed out in Hammond, 1971.) Even though a small number of subjects was employed ($N = 5$), the experiment yielded clear results. Performance (r_a) reached the limit of achievement after two or three comparisons. It

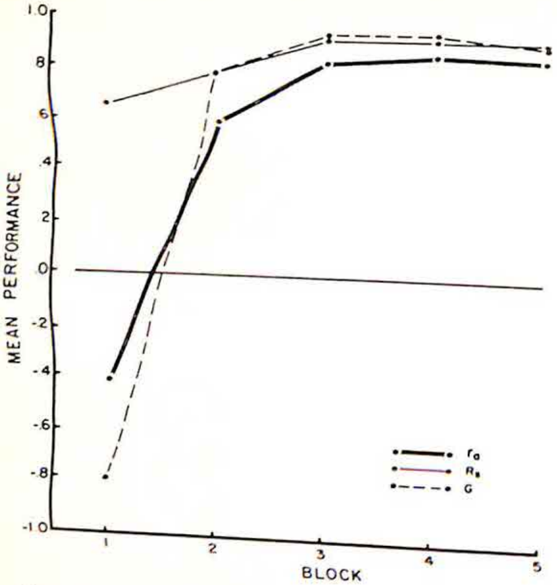


FIG. 6. Indexes of achievement (r_a), knowledge (G), and control (R_s) in a nonlinear inference task when cognitive feedback is presented in the form of graphic displays. (Block = 10 trials.)

should be noted that the acquisition of knowledge (G) by these subjects did not differ from that observed under outcome feedback conditions in the Deane et al. study; that is, G was near optimal in both experiments. When subjects received *only* graphics feedback, however, control (R_s) *as well as* G was high (see Figure 6). In short, feedback which provides the appropriate cognitive material permits cognitive control to be executed and overall performance to improve.

SUMMARY

Our purpose has been to demonstrate that cognitive control is a concept of theoretical significance, and that it is susceptible to quantification and experimental manipulation. In addition, when this concept is brought to bear on important cognitive matters—individual learning, clinical judgment, and interpersonal conflict—it leads to explanations of behavior which are meaningful and coherent across all three domains.

Furthermore, the evidence which suggests that traditional, response-oriented outcome feedback is an impediment to cognitive control (and thus to performance) also points to the facilitating effect of *cognitive* material as feedback. This shift in conception of the notion of feedback carries considerable practical as well as theoretical significance, for it is now evident that computer technology can be used to produce such facilitating feedback. Therefore, it should be possible to improve performance in complex cognitive tasks in which improvement has heretofore been unlikely.

REFERENCES

AZUMA, H., & CRONBACH, L. J. Cue-response correlations in the attainment of a scalar concept. *American Journal of Psychology*, 1966, **79**, 38–49.

BJORKMAN, M. Studies in predictive behavior: Explorations into predictive judgments based on functional learning and defined by estimation, categorization, and choice. *Scandinavian Journal of Psychology*, 1965, **6**, 129–156.

BREHMER, B. Cognitive dependence on additive and configural cue-criterion relations. *American Journal of Psychology*, 1969, **4**, 490–503.

BREHMER, B., AZUMA, H., HAMMOND, K. R., KOSTRON, L., & VARONOS, D. A cross-national comparison of cognitive conflict. *Journal of Cross-Cultural Psychology*, 1970, **1**, 5–20.

BREHMER, B., & LINDBERG, L. Retention of probabilistic cue-criterion relations as a function of cue validity and retention interval. *Journal of Experimental Psychology*, 1970, **86**, 331–334.

BRUNSWIK, E. *Perception and the representative design of psychological experiments.* Berkeley: University of California Press, 1956.

CASTELLAN, N. J. *Comments on the "lens model" equation and the analysis of multiple cue judgment tasks.* (Indiana Mathematical Psychology Program Rep. No. 71–1). Bloomington: Indiana University, 1971.

DAWES, R. A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist*, 1971, **26**, 180–188.

DEANE, D. H., HAMMOND, K. R., & SUMMERS, D. A. Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology*, 1972, **92**, 20–26.

DEESE, J., & HULSE, S. H. *The psychology of learning.* New York: McGraw-Hill, 1967.

DUDYCHA, L. W., & NAYLOR, J. C. Characteristics of the human inference process in complex choice behavior situations. *Organizational Behavior and Human Performance*, 1966, **1**, 110–128.

GOLDBERG, L. R. Simple models or simple processes? Some research on clinical judgments. *American Psychologist*, 1968, **23**, 483–496.

GOLDBERG, L. R. Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. *Psychological Bulletin*, 1970, **73**, 422–432.

HAMMOND, K. R. Probabilistic functionalism: Egon Brunswik's integration of the history, theory and method of psychology. In K. R. Hammond (Ed.), *The psychology of Egon Brunswik.* New York: Holt, Rinehart & Winston, 1966.

HAMMOND, K. R. Computer graphics as an aid to learning. *Science*, 1971, **172**, 903–908.

HAMMOND, K. R., & BREHMER, B. Quasi-rationality and distrust: Implications for international conflict. In L. Rappoport & D. A. Summers (Eds.), *Human judgment and social interaction.* New York: Holt Rinehart & Winston, in press.

HAMMOND, K. R., HURSH, C., & TODD, F. J. Analyzing the components of clinical inference. *Psychological Review*, 1964, **71**, 438–456.

HAMMOND, K. R., & SUMMERS, D. A. Cognitive dependence on linear and nonlinear cues. *Psychological Review*, 1965, **72**, 215–224.

HAMMOND, K. R., WILKINS, M., & TODD, F. J. A research paradigm for the study of interpersonal learning. *Psychological Bulletin*, 1966, **65**, 221–232.

HILGARD, E. R. *Theories of learning.* (2nd ed.) New York: Appleton-Century-Crofts, 1956.

HURSCH, C., HAMMOND, K. R., & HURSCH, J. Some methodological considerations in multiple-cue probability studies. *Psychological Review*, 1964, **71**, 42–60.

KNOWLES, B. A., HAMMOND, K. R., STEWART, T. R., & SUMMERS, D. A. Positive and negative redundancy in multiple-cue probability tasks. *Journal of Experimental Psychology*, 1971, **90**, 157–159.

LEEPER, R. Cognitive learning theory. In M. H. Marx (Ed.), *Learning: Theories.* New York: Macmillan, 1970.

MANDLER, G. Verbal learning. In G. Mandler, P. Mussen, N. Kogan, & M. A. Wallach (Eds.), *New directions in psychology: III.* New York: Holt, Rinehart & Winston, 1967.

NAYLOR, J. C., & CLARK, R. D. Intuitive inference strategies in interval learning tasks as a function of validity magnitude and sign. *Organizational Behavior and Human Performance*, 1968, **3**, 378–399.

ROZEBOOM, W. Comments on Professor Hammond's paper. In J. Royce & W. Rozeboom (Eds.), *Psychology of knowing.* London: Gordon & Breach, 1971.

SHEETS, C., & MILLER, M. J. The effect of cue-criterion function form on multiple-cue probability learning. *American Journal of Psychology*, in press.

SUMMERS, D. A. Adaptation to change in multiple probability tasks. *American Journal of Psychology*, 1969, **82**, 235–240.

SUMMERS, D. A., & HAMMOND, K. R. Inference behavior in multiple-cue tasks involving both linear and nonlinear relations. *Journal of Experimental Psychology*, 1966, **71**, 751–757.

TODD, F. J., & HAMMOND, K. R. Differential feedback in two multiple-cue probability learning tasks. *Behavioral Science*, 1965, **10**, 429–435.

TUCKER, L. R. A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review*, 1964, **71**, 528–530.

UHL, C. Learning interval concepts: I. Effects of differences in stimulus weights. *Journal of Experimental Psychology*, 1963, **66**, 264–273.

WARD, J. H., & DAVIS, K. Teaching a digital computer to assist in making decisions. In L. J. Cronbach & G. C. Gleser (Eds.), *Psychological tests and personnel decisions.* Urbana: University of Illinois, 1965.

(Received April 22, 1971)

Psychological Review
1972, Vol. 79, No. 1, 68–92

TIMING AND DISCRIMINATION OF SHOCK DENSITY IN AVOIDANCE[1]

JOHN GIBBON[2]

New York State Psychiatric Institute

A theoretical account of asymptotic avoidance behavior is proposed which deals with shock density effects. Behavior at asymptote is characterized by the operation of two processes. An elementary scalar timing process generates response latencies proportional to the length of the preshock interval. However, responding occurs only when subjects discriminate a "worthwhile" or biased difference between intershock intervals containing at least one response, and intershock intervals containing no responses. The two processes operating in avoidance schedules result in stochastic systems known as semi-Markov chains. Analysis of these chains are shown to produce first-order descriptions of response rates and shock rates on the standard uncued free operant avoidance schedule and on several variants. The formulations for the discrimination process require a strong bias in favor of not responding. Extrapolation to classical cued avoidance suggests that a similar bias is responsible for the "avoidance decrement" phenomenon. A less extreme bias may account for maintained responding on a retractable lever in a discrete trial procedure in which responses postpone but do not eliminate shock.

Early work in free operant avoidance (Sidman, 1953, 1954) and later reviews (Anger, 1963; Sidman, 1966) have argued for the importance of temporal discriminations of shock-due time as a critical determinant of avoidance performance. A recent quantitative account (Gibbon, 1971) has shown that much of the parametric data in avoidance is well described by an elementary scalar timing process underlying temporal discriminations in avoidance latencies. The theory holds that once the response has been learned, response strength does not change, so that changes in avoidance rate are solely a function of changes in timing. The scope of the account, however, has been limited to situations in which the typical avoidance performance results in a relatively low frequency of shock.

A growing literature exists in which shock frequency per se, independently of timing behavior, appears to be an important determinant of avoidance (Herrnstein, 1969; Herrnstein & Hineline, 1966; Sidman, 1962). In these experiments, avoidance behavior of rats has been shown to be sensitive to the *amount* of the reduction in shock frequency produced by responding, relative to shock frequency when subjects do not respond at all. The data have been advanced, however, more as a challenge to existing accounts of avoidance (e.g., two-factor theory), than as evidence for an alternative. Thus, while it is clear from these experiments that reduction of shock density is an important variable in avoidance, it is not at all clear how it exerts control.

This paper proposes that shock density effects are mediated by a discrimination process which is sensitive to the relative improvement in shock frequency produced by the scalar timing performance. Timing is discussed first, and then the scalar assumption is used in elaborating the discrimination process for a variety of avoidance schedules. Timing in these schedules results in stochastic processes called semi-Markov chains, and the discrimination mechanism may be incorporated into the chains in a simple way. The analysis is applied to three free operant paradigms: standard uncued avoidance (Sidman, 1953), uncued avoidance when shocks are omitted with some probability (Boren & Sidman, 1957), and avoidance when shocks are delivered with different probabilities after responding and not responding (Herrnstein & Hineline, 1966). Two discrete trial situations are then studied: classical cued avoidance with a wheel-turning response (avoidance decrement), and "avoidance" with a retractable lever response which postpones but does not eliminate shock (Hineline, 1970).

---

[1] Reported in part in a paper titled "An Elastic Stochastic Timer and Detection in Free Operant Avoidance" read at the Mathematical Psychology meetings, Ann Arbor, August 1969. Preparation of the manuscript was facilitated by National Institutes of Health Grants GRS-5-SO-1-FRO-5650 and MH19082 to J. Gibbon, and MH-07279 to H. F. Hunt. I wish to thank A. Clymer and J. Horn for assistance in data analysis. I am also particularly grateful to G. G. Neffinger for permission to include data from an unpublished study.

[2] Requests for reprints should be sent to John Gibbon, New York State Psychiatric Institute, 722 West 168th Street, New York, New York 10032.

TIMING

Avoidance procedures define an interval of time, the preshock interval, within which a prescribed response is effective in postponing or eliminating shock. Shock is delivered whenever the preshock interval elapses without a response. In classical cued avoidance, the preshock interval begins with the onset of a warning signal, while in uncued free operant avoidance, preshock intervals begin with each response and shock.

Subjects are usually thought of as "timing" the preshock interval when responding tends to occur late in the interval, shortly before the time at which shock would otherwise be delivered. As Anger (1963) and others have pointed out, however, these later portions of the interval occur less frequently than early portions, since each response occurrence precludes an opportunity to respond later in the interval. A measure of timing behavior which corrects for this difficulty is the conditional probability of a response at time t given a wait of at least t, that is, given an opportunity to respond at time t. When this measure rises as the end of the interval is approached, subjects are thought to be timing the interval even when short latencies are more frequent than long ones. When conditional probability is flat over the preshock interval (indicating an exponential latency distribution) or is decreasing, subjects are viewed as not timing.

This view of timing seems unreasonable on two counts. First, the use of an exponential or steeper drop in probability density as a criterion is arbitrary. Any number of equally "random" latency mechanisms generate rising conditional density functions. For example, a latency which is composed of several subresponses, each of which is exponential, results in a rising conditional density function. Yet each subresponse, regarded singly, meets the criterion for no temporal discrimination.

A second objection is more fundamental. It rests on the simple requirement that timing behavior be tied to the size of the timing interval, so that changes in shock-due time produce corresponding changes in behavior. A strict construction of this view would hold that no conclusion whatever about timing may be drawn from observation of a single latency distribution. Rather, timing is indicated by changes in latency with changes in the preshock interval. The scalar timing process espoused here is just one, and possibly the simplest, of many ways in which latencies might change with changes in shock-due time.

In Figure 1, hypothetical latency distributions corresponding to no timing and to two different kinds of timing have been sketched. The top panel shows a probability density for response latencies with a preshock interval of one unit of time. The tail of the distribution (stippled area to the right of 1.0) represents latency values longer than the preshock interval which might be observed if shocks were occasionally omitted. The area in the tail (q) is shock probability. The lower three panels represent distributions corresponding to a preshock interval of two units of time. The first of these, labeled No Timing, represents a complete lack of correlation with the changed shock delivery time. The new distribution is identical to the original, and shock probability (q') has been reduced by an amount equal to the area under the curve between

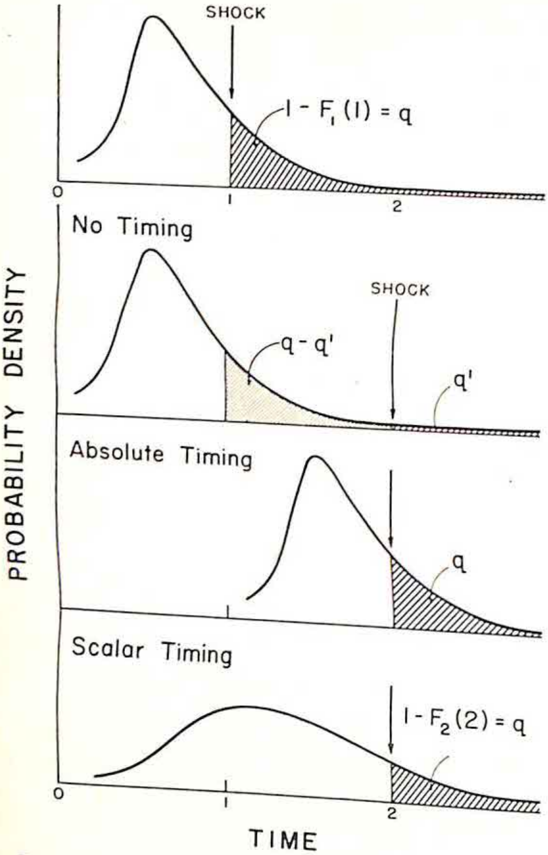


FIG. 1. Probability densities for a preshock duration of one unit of time (top panel), and for two units of time (lower three panels). (The lower three panels represent, respectively, no change from the unit distribution (No Timing), an absolute shift of one unit of time (Absolute Timing), and a relative shift of one unit of time (Scalar Timing).)

1.0 and 2.0. The third panel, Absolute Timing, represents latencies which are tied to the amount of real time before shock delivery. The distribution is thus shifted to the right by an additive constant equal to the increment in shock-due time. Latencies in the 0.0 to 1.0 interval here have an ambiguous status, since latencies at this remove from shock-due time are not observable in the unit case. In practice, timing of this sort would be indicated if the form of the distribution between 1.0 and 2.0 were similar to the form of the unit timing distribution between 0.0 and 1.0, and if shock probability were close to the original.

The bottom panel represents an example of the scalar timing process proposed here as the basic latency mechanism in avoidance. The unit distribution is simply scale transformed for different values of shock-due time. In this example, the time axis is "stretched" by a factor of 2, so that probability density is reduced at corresponding time values, but shock probability (q) remains the same.

In the scalar process, the time unit is arbitrary. *For example, if the bottom panel represented the unit timer, then the top panel would represent this timing distribution rescaled by* $\frac{1}{2}$. The scalar property translates directly into expected values,

$$\mu_r = r\mu \quad [1a]$$

$$\sigma_r = r\sigma, \quad [1b]$$

where μ and σ are the mean and standard deviation of the unit latency distribution, and r is the preshock duration.[3]

The scalar process is intended as an idealized description of asymptotic avoidance latencies. If the description is taken literally, several complications immediately arise. For example, at short preshock durations, the time required to physically execute the response (motor time) may represent a significant portion of the recorded latency. Variability as well as time is then added to the scalar component, and the resulting latencies are not scalar for short preshock durations. The size of the motor component, however, is small relative to the timing component for moderate preshock intervals.

Very long preshock intervals pose a difficulty also. The "Weber-like" variance property (Equation 1b) suggests that scalar timing may break down at some

---

[3] Equations 1a and 1b hold for scalar distributions like those sketched in Figure 1, and also for these distributions truncated at shock delivery. The scalar property for the truncated distributions is

$$\frac{f_r(t)}{F_r(r)} = \frac{f(t/r)}{rF(1)}, \quad 0 \le t < r,$$

where f_r, F_r, and f, F are the densities and distribution functions for timing distributions corresponding to preshock intervals of r and unity, respectively. Latencies reported in the literature are perforce truncated at shock-due time, and Equations 1a and 1b are to be interpreted as applying to the truncated distributions.

upper limit, analogous to the upper limit for Weber's Law in sensory psychophysics. While significant deviations do occasionally appear at very long preshock durations, the scalar property holds approximately, for the data considered here, over a broad range of intermediate values.

Another complicating factor is the well-established finding of dependencies between successive latencies (Sheffield, 1948; Wertheim, 1965) between latencies obtained early and late in a single session (Hineline & Herrnstein, 1970; Wertheim, 1965), and between successive sessions, even after extended training (Hineline & Herrnstein, 1970). Convincing evidence for scalar timing requires pooling over several sessions and even several determinations at each preshock duration. The pooled distributions are then mixtures of heterogeneous component distributions. It may be shown (Gibbon, 1971) that mixtures of this sort are scalar if the component distributions are scalar. It thus appears that a complex system of dependencies between successive avoidance responses and shocks may nevertheless exhibit a very simple dependence on the length of the preshock interval, namely, the scalar property.

DISCRIMINATION OF SHOCK DENSITY

The shock density effects to be described have a common relativistic feature. Response strength is directly related to the difference between two rates: the shock rate when subjects are working at the avoidance task, and the shock rate when they are not responding at all. When this difference is sufficiently small, responding is eliminated. When both shock rates are high, the effect is "punishmentlike." When both shock rates are close to zero, the effect is "extinctionlike." The discrimination process proposed here holds that responding declines in both cases for the same reason—that the two rates are difficult to tell apart.

Loosely speaking, subjects are viewed as discriminating whether or not their timing performance results in a "worthwhile" improvement in shock density. Under some experimental arrangements, the improvement is very substantial and avoidance responding is reasonably well described by scalar timing alone. Under other experimental arrangements, an improvement is difficult to discriminate, and avoidance behavior shows a corresponding decline. The standard uncued free operant avoidance situation provides examples of both sorts of behavior. The discrimination process will be outlined for this situation first, and then applied to more complex experimental paradigms.

Standard Uncued Avoidance

In the standard uncued avoidance schedule, shocks are programmed at a constant time, r seconds, from a preceding response, and a possibly different constant time, s seconds, from a preceding shock when no response intervenes. Responses recycle the postresponse apparatus clock (r-clock), so that all shocks would be avoided if subjects always responded within r seconds of the preceding response. For $r > s$, a well-trained rat may receive as few as 1 out of every 100 shocks which would be delivered by the s-clock if he did not respond. However, when r is small relative to s, even very good avoidance performances occasionally result in intershock intervals which are shorter than s. Sidman (1953) likened the result to a punishment effect, since shocks which closely follow a response appear to suppress subsequent behavior, so that as r is made progressively smaller, responding is eliminated. In the context of a discrimination of shock density differences, this effect will be seen to depend on the value of s as well, since responding at a small r value may still produce substantial improvement in shock density relative to the shock rate on the s-clock when s is also small.

The realization of the discrimination process is a straightforward adaptation of detectability theory. After each shock, the rat samples two time values. One is s, the time between shocks without a response, and the other is the value assumed by a random variable, x, which is the time between two shocks including at least one

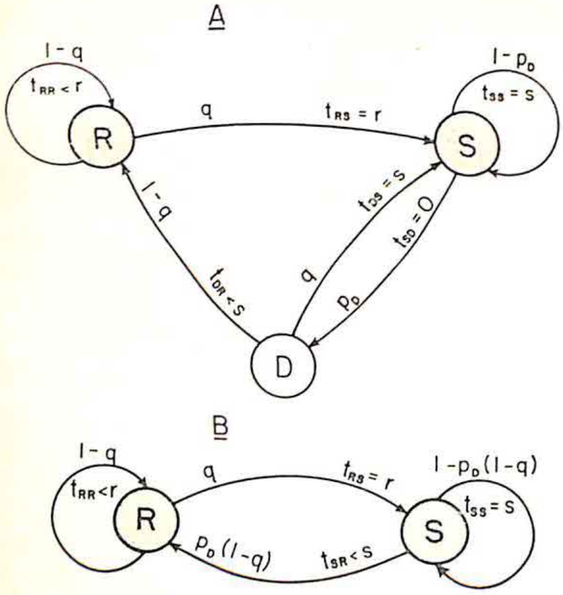


FIG. 2. State diagrams for chains corresponding to the standard uncued free operant avoidance schedule. (The three state chain (A) represents the discrimination process. When subjects discriminate an improvement in shock density produced by responding, they enter State D before continuing to work on the avoidance schedule. For the two-state semi-Markov chain (B), the D state has been incorporated into the S state and transition probabilities and times adjusted accordingly.)

intervening response. The two values are compared, and discrimination of a worthwhile improvement occurs when x is sufficiently larger than s. The comparison is stated in terms of time rather than rate for simplicity. For example, if n intershock intervals, x_i, were sampled, the comparison of shock rates would ask whether or not $\sum x_i > ns$.

The intershock interval comparison is made according to a biased decision rule. Discrimination of a worthwhile difference between x and s occurs whenever

$$bx > s. \qquad [2]$$

The constant, b, is a multiplicative input bias which systematically distorts the real improvement in intershock intervals produced by a subject's avoidance performance. Generally, x must be considerably larger than s for discrimination of an improvement to occur. When the criterion, 2, is not met, subjects are assumed to simply wait for the next shock, when a new value of x is sampled and compared with s. When the criterion is met, subjects attempt the avoidance task, with response latencies governed by scalar timing.

These assumptions are described by the state chain in Figure 2A. In the response state (State R), subjects recycle the r-clock with variable interresponse times ($t_{RR} < r$) governed by scalar timing. These avoidance responses occur with probability $1 - q$ (Figure 1). When r seconds elapse without a response ($t_{RS} = r$), a shock is delivered (State S), and this transition occurs with probability q. After each shock, a discrimination trial occurs, and the criterion, 2, is met with some probability, p_D. Subjects are then thought of as entering a "discriminate" state (State D) in which avoidance latencies are again controlled by scalar timing. Discrimination trials are assumed to take negligible time ($t_{SD} = 0$), and timing is assumed to be scalar across the r- and s-clocks. Thus after a successful discrimination, the first response is distributed as the scalar timer for s and occurs with probability $1 - q$.

Procedurally, the system in Figure 2A is a discrete trials procedure with no intertrial interval, in which trials begin and end with responses and shocks, and have variable durations. The three-state chain may be reduced to a two-state chain by combining the S and D states. This results in the mathematically more tractable semi-Markov chain in Figure 2B. Such chains are like discrete Markov chains in that transition probabilities are constant. However, transition times may depend on the state to be entered as well as the state occupied, and the system is not, in general, Markov in time.[4]

In the context of Figure 2, response and shock *rates* are the reciprocals of the mean recurrence times for the R and S states. Mean recurrence time formulations for semi-Markov chains are generalizations of similar expressions for discrete chains, and those results of earlier work which are required for the analysis of interresponse and intershock times have been summarized in Appendix A. Before using these results, a functional relationship is required between discrimination probability, p_D, and the settings of the apparatus clocks, r and s.

---

[4] A chain which is Markov in time has constant (instantaneous) transition probabilities conditional at most on the state occupied. This results in exponentially distributed transition times (see, e.g., Feller, 1966, p. 316).

Discrimination probability. The discrimination process as outlined above requires a comparison between sample values of the random variable x, and s. From Figure 2B, the time between shocks when there is at least one intervening response is

$$x = t_{SR} + \tau + r, \qquad [3]$$

where τ is the sojourn time in State R excluding the final transition to State S. τ is a mixture or "randomly stopped sum" with parameter n for the number of responses in the r-clock before the first shock.

$$\tau = \sum_{i=0}^{n} t_{RR_i}, \qquad [4]$$

where, if responses are independent,

$$P(n = k) = (1 - q)^k q, \quad k = 0,1,\cdots, \qquad [5]$$

and q is shock probability (Figure 1). An exact solution for discrimination probability according to the biased decision rule, 2, requires a knowledge of the distribution of τ. For an efficient avoidance performance, a symmetric form for τ is not realistic. For a good avoider, q is small and the geometric mixing distribution (Equation 5) dominates τ, producing considerable skew in the distribution of x. In Appendix B an approximate solution is obtained assuming an exponential form for x. This approximation is substantively equivalent to assuming that subjects disregard variability in the transition times of Figure 2, but attend to their means.

The solution shows p_D to be a function of the *ratio* of r and s,

$$p_D = \begin{cases} \exp\left[\dfrac{-q}{(1-q)\mu}\left(\dfrac{s}{r}\left[\dfrac{1}{b} - \mu\right] - 1\right)\right], & \dfrac{r}{s} < \dfrac{1}{b} - \mu \\ 1.0 \quad \text{otherwise,} & \end{cases} \qquad [6]$$

where μ is the mean of the unit latency distribution, q is shock probability, and b is the bias parameter. The dependence of p_D on the ratio of r and s (rather than, e.g., the difference) is an important feature of Equation 6 which stems from the scalar expectations for the latency distributions. The qualitative observation that response strength declines as r is made small relative to s has an intuitive counterpart in the above solution, since $p_D \to 0$ as $(r/s) \to 0$.

The dependence of p_D on avoidance performance and bias is also appropriate in view of the parameter interpretations. Discrimination probability approaches zero whenever (*a*) avoidance performance is very inefficient ($q \to 1$), (*b*) latencies are so short as to add very little time to the interval between shocks ($\mu \to 0$), or (*c*) bias heavily favors not responding ($b \to 0$).

Evidence on the feasibility of this solution for the standard uncued avoidance paradigm is provided by response and shock rates reported in Sidman's original investigation. From Appendix A, the mean recurrence time formulation (Equations A6a and A6b) for the two-state chain in Figure 2B may be normalized for s to provide interresponse and intershock times as functions of r/s.

$$\frac{m_{RR}}{s} = \frac{r}{s}\left[\mu + q(1-\mu)\right] + \frac{q}{p_D(1-q)} + q(1-\mu) \qquad [7a]$$

$$\frac{m_{SS}}{s} = \left[\frac{p_D(1-q)}{q}\right]\frac{m_{RR}}{s}, \qquad [7b]$$

where m_{RR} and m_{SS} are mean interresponse and intershock times, and p_D is given by Equation 6. Equations 7a and 7b are constant for constant r/s; thus, data collected at different absolute r, s pairs with the same ratio should superimpose. In Figure 3, data from one of Sidman's original subjects have been replotted on double log coordinates. The curves through the data are plots of Equations 7a and 7b. The parameter values were obtained by first fitting a linear function by eye to the response data at large r/s values, and then adjusting the bias parameter so that the response and shock curves would intersect at $r/s = .71$. This resulted in a bias

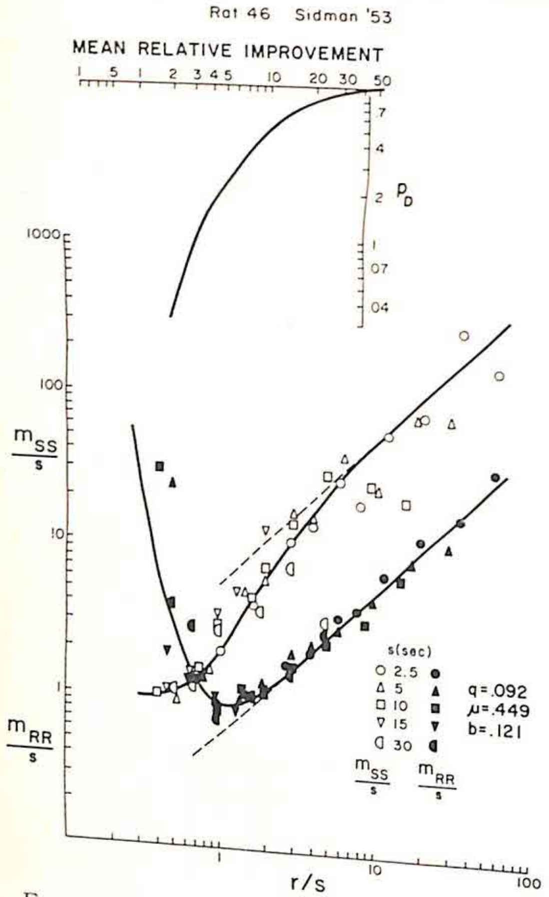


FIG. 3. Mean interresponse time (filled points) and intershock time (open points) as a function of r/s for one subject. (Double-log coordinates have been used to encompass the broad range of the functions, and also because they visually compensate for increasing variances with means, a property of scalar timing.)

heavily in favor of not responding. The agreement between data and theory is impressive for the mean interresponse time function. The shock data are more variable, but they are based on a smaller sample. The discrimination probability function for these parameter values is shown above the data curves. The p_D ordinate has been set at the r/s value at which p_D reaches 1. The upper abscissa axis, labeled "mean relative improvement" represents $E[(x-s)/s]$, the expected relative improvement in intershock intervals containing avoidance responses. Mean relative improvement is linear in r/s, and has been scaled here to emphasize the degree to which bias favors not responding. Improvement is better than an order of magnitude before p_D equals .6, and p_D is virtually zero when improvement is still better than 100%. The dashed curves are the response and shock predictions without the discrimination process (i.e., assuming $p_D = 1$). The effect of the drop in discrimination probability with decreasing r/s is discernible earlier in the shock function than in the response function. This is because response–response latencies, when they occur, decrease proportionally with decreasing r/s values according to the scalar rule. The exponential drop in p_D, however, eventually dominates the system as entrances into the r-clock become less and less frequent. Mean intershock time, on the other hand, decreases directly with decreasing discrimination probability, and multiplying out Equation 7b, m_{SS}/s is seen to approach 1 as r/s approaches 0.

Two other subjects were run in the original study and their data, while more variable, show the same basic features evident in Figure 3. In particular, the bias in favor of not responding may be shown to be at least as strong as the bias for rat 46. In a loose, anthropomorphic sense, subjects appear to undervalue the real improvement they are able to produce in intershock intervals by working.[5] A strong bias in favor of not responding will reappear later where it will be seen to bear an important relation to asymptotic extinction.

---

[5] The strength of the bias in favor of not responding motivated the adoption of a multiplicative rather than an additive decision rule. For example, the minimum interresponse time for all three subjects occurs at an r/s value of approximately 1.0. For equal clock settings, there is always some improvement in the time between shocks when at least one response intervenes, even if *only* one response intervenes. Yet p_D values required to fit these subjects' data at $r/s = 1.0$ are in the neighborhood of .2, and the mean intershock time for working at these clock settings is approximately six times the value of s. A low-variance Gausian discrimination mechanism with distributions centered on s and $E(x)$, producing a sufficient steep psychometric function to fit the data, would require an additive criterion of better than seven σ units away from s. In the light of this amount of strain for a more conventional discrimination mechanism, the multiplicative bias seemed intuitively more appealing.

The theoretical curves in Figure 3 are first approximations to the data rather than definitive descriptions. There is considerable variability around the shock function and a suggestion of some systematic effects due to the *s*-clock setting which are ignored in this formulation. However, the theory is certainly a parsimonious description (27 data points per parameter value), and the superimposition of the data at constant r/s ratios represents important evidence for the scalar rule and for a discrimination mechanism sensitive to relative rather than absolute r, s values.

In the standard uncued avoidance paradigm, responding is eliminated as intershock intervals obtained when working at the avoidance task approach the time between shocks when no responses are made. Sidman originally described the effect in terms of a "punishment gradient," with shocks increasing in punishment power as the response–punishment interval is decreased. While the idea of a punishment gradient remains imprecise, it is intuitively appealing in its simplicity and in its direct relation to the well-known efficacy of traditional punishment procedures in eliminating behavior. The role of the *s* interval in a punishment interpretation is not immediately apparent. However, one might regard shorter *s* intervals as establishing higher avoidance motivation levels, or possibly as establishing a "punishment gradient for not responding." The latter possibilities are qualitatively close to the present discrimination proposal, since a comparison is required between the consequences of responding and not responding.

Shock Probability in Uncued Avoidance

A traditional alternative to punishment for eliminating behavior is extinction. While the results of the two procedures are the same, the underlying mechanisms are usually assumed to be quite different. Punishment is thought of as *suppressing* behavior, or motivating *not* responding, while the traditional avoidance extinction procedure in which shock is simply discontinued is thought to exert its effect by removing the motivation *for* responding.

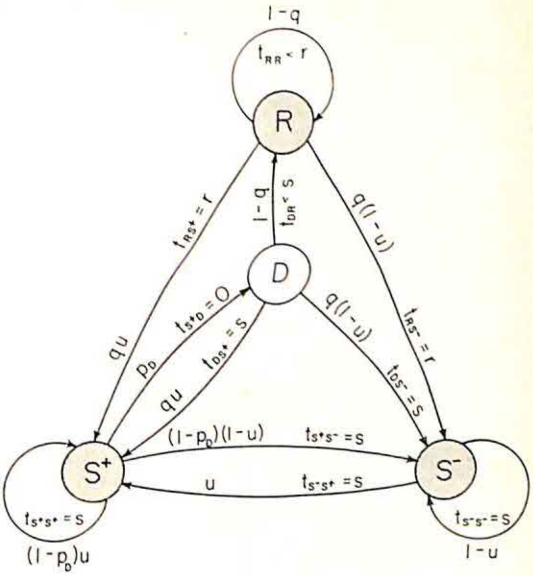


FIG. 4. State diagram for the shock probability experiment.

Extinction conditions may be approached parametrically in the standard uncued avoidance paradigm by reducing the probability with which programmed shocks are delivered. When the discrimination process is applied to this situation, extinction may be viewed as the endpoint on a continuum along which intershock intervals for responding *and* not responding are both increasing and becoming more variable. The result is that discrimination of an improvement in the response distribution occurs less and less frequently as shock probability approaches zero.

Boren and Sidman (1957) studied the shock probability variable and found that responding was only slightly reduced by lowered shock probability, until shock probability was dropped to very low values. In terms of the discrimination analysis, this finding will be seen to depend on the level of the original avoidance performance, and on the choice of the r and s parameter values.

In the Boren and Sidman experiment, shocks which became due in either the *r*- or *s*-clock met the same probability, u, of being delivered to the subject. Whenever a shock became due, the *s*-clock was recycled whether or not the shock was delivered. In Figure 4, the state chain

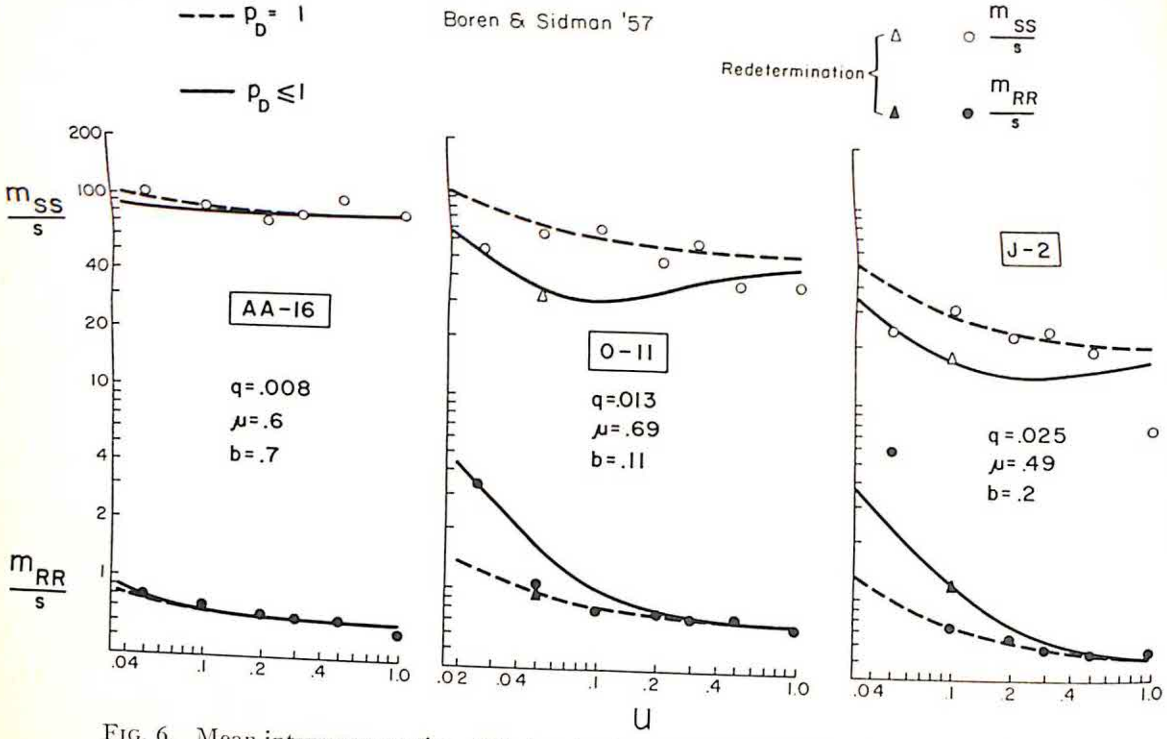


FIG. 6. Mean interresponse time (filled points) and intershock time (open points) as a function of shock probability (u) for three rats.

two subjects had larger shock rates, indicating higher qs, and it was felt that for these subjects, the assumption of negligible responding in S^- was not tenable. The dashed functions are predictions assuming perfect discrimination ($p_D = 1$), and the solid curves assume the discrimination process with the indicated bias. For all subjects, omission of shocks produces small, though discernible, changes in responding until u is very small. For one subject, AA-16, responding is only slightly attenuated even when only 5% of the due shocks are delivered. According to the theory, this result depends on a highly efficient avoidance performance. Examination of Equation 11 reveals that discrimination probability increases with decreasing q. That is, efficient avoidance behavior results in a more discriminable improvement in intershock intervals. As shock probability is lowered, intershock times are lengthened *both* for working and not working; with the result that discriminability is not substantially impaired until very low values of u are studied, which produce considerable overlap in the two distributions (Figure 5). In Figure 6, subjects are presented in the order of increasing qs (left to right), and it can be seen that the solid and dotted response curves begin to diverge at progressively larger values of u.

For two of the subjects, the response data diverge more rapidly than the discrimination process predicts. These data would appear to require a steeper p_D function for a more realistic description. Since the approximations involved act to reduce variance in the intershock interval distributions, further reduction in variance may require an averaging assumption.

The theory appears to provide a reasonable first approximation to the data, with appropriate qualitative distinctions between different levels of efficiency in the avoidance performance. Important confirmatory or disconfirmatory evidence would be provided by either of two variants of the experimental paradigm. If shocks which became due on the r-clock were always delivered, but shocks in the s-clock were delivered with probability u, the theory predicts that response rate would begin to decline at much larger values of u.

An alternative arrangement, which is more closely related to the previous application, would be to run the identical experiment, but vary the r/s ratio. Small r/s values should produce an early decline in responding with decreasing u, while large r/s values should make responding resistant to change with decreasing u.

Differential Shock Probability for Responding and Not Responding

In the shock probability experiment, shocks were omitted with equal frequency in the r- and s-clock. An important generalization of this paradigm, which combines extinctionlike and punishmentlike features, has been devised by Herrnstein and Hineline (1966). In their procedure, only the first response after a shock has experimental consequences, and shock probability after this response is programmed independently from shock probability after a preceding shock. A state chain diagram for this experiment is shown in Figure 7. The first response after a shock initiates a response clock (State R) set at r seconds. This clock does *not* recycle with subsequent responses. At the end of the r cycle, a shock is delivered (State S^+) with probability u_1, or not delivered (State S_R^-) with probability $1 - u_1$. When shock is not delivered, the r-clock recycles, meeting the same shock probability until the first postresponse shock is delivered. The response clock is then stopped, and the shock clock set at s seconds is initiated. Shocks are assumed to be followed by discrimination trials as before, with discrimination of an improvement occurring with probability p_D. When a discrimination trial favors responding, subjects attempt to time the s interval and respond in time $(t_{S^+R} < s)$ with probability $1 - q$. The discriminate state (State D) has been incorporated with the shock state, so that the $S^+ \rightarrow R$ transition occurs with probability $p_D(1 - q)$. When responses do not occur, the s-clock times out and meets a probability of shock, u_0, which is set independently of u_1. Shocks therefore follow one another with probability $u_0[1 - p_D(1 - q)]$. When shock is not delivered at the end of the s-clock interval, the clock recycles (State S_S^-) meeting the same shock probability, u_0, until the next shock occurs. As in the previous analysis, it is assumed that after a due shock has not been delivered (S_S^-), responding is negligible. This assumption is possibly less reasonable here than in the preceding case, since the clock settings in this experiment were small ($r = s = 2$ seconds). However, responding was reported to be most frequent immediately after shock, and the data reported do not have the fine grain required to discriminate alternatives.

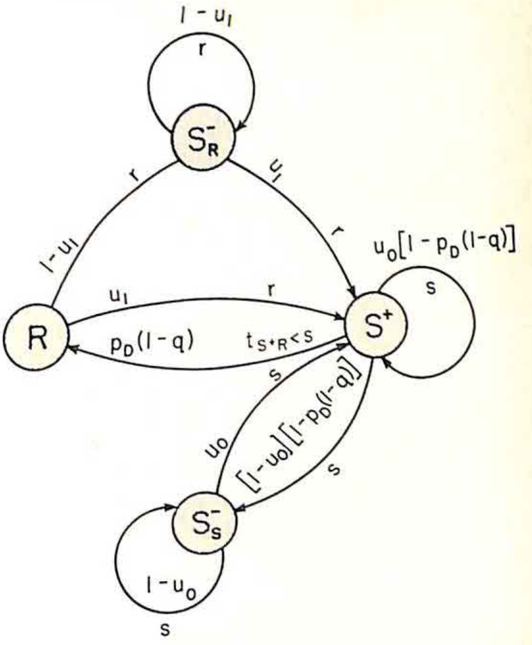


FIG. 7. State diagram for the differential shock probability experiment.

Discrimination probability. When responses do not occur, shocks are geometrically distributed at multiples of s seconds,

$$x_0 = (n_0 + 1)s. \qquad [14]$$

Similarly, intershock intervals including a response are distributed as

$$x_1 = t_{DR} + (n_1 + 1)r \qquad [15]$$

where

$$P(n_i = k) = (1 - u_i)^k u_i, \quad k = 0, 1, \cdots,$$

for $i = 0, 1$.

In Appendix D, an approximate solution for the discrimination problem is developed which turns out to be a special case of the preceding analysis. The result is

$$p_D = \begin{cases} \dfrac{\lambda_0}{\lambda_0 + \lambda_1} e^{-\lambda_1 K}, & K \geqslant 0 \\ 1 - \dfrac{\lambda_1}{\lambda_0 + \lambda_1} e^{\lambda_0 K}, & K < 0, \end{cases} \qquad [16]$$

where

$$\lambda_0 = \frac{bu_0}{1 - u_0}, \quad \lambda_1 = \frac{u_1}{1 - u_1},$$

and

$$K = \frac{1}{b} - \mu - 1.$$

Herrnstein and Hineline ran several values of u_1 with u_0 fixed, and several values of u_0 with u_1 fixed. In the former case, when the shock schedule for not responding was fixed, response strength increased as shock probability after a response was decreased. For the discrimination process, this corresponds to the fact that in Equation 16,

$$\lim_{u_1 \to 0} p_D = 1. \qquad [17]$$

The effect is a straightforward analog of increasing the *r*/*s* ratio by increasing *r* in the uncued avoidance paradigm. As post-response shock probability is made small, more time is spent in the $S_{\bar{R}}$ state (Figure 7) and improvement in intershock intervals containing a response becomes easier to discriminate.

The data also show that when the post-response shock schedule (u_1) was fixed, and the postshock shock schedule was varied, response strength increased directly with an increase in u_0. In the discrimination process (Equation 16), this finding corresponds to

$$\lim_{u_0 \to 1} p_D = \begin{cases} \exp\left(\dfrac{-u_1 k}{1 - u_1}\right), & K \geqslant 0 \\ 1, & K < 0. \end{cases} \qquad [18]$$

Discrimination probability increases as shocks for not responding are delivered more frequently. If bias favors responding ($K < 0$), discrimination is perfect when shock at the end of the *s*-clock is certain. However, when bias sufficiently favors not responding ($K > 0$), discrimination of an improvement in intershock intervals is less than perfect. The effect might be viewed as a sort of "punishment of not responding" result, analogous to increasing the *r*/*s* ratio by decreasing *s* in the uncued avoidance situation.

An important focus of the experiment was the demonstration of a generalized extinction procedure. The authors found that responding could be extinguished by simply setting the two shock probabilities *equal* irrespective of their absolute value. Letting $u_1 = u_0 \equiv u \neq 0, 1$ in Equation 16,

$$p_D = \frac{b}{1 + b} \exp\left(\frac{-uK}{1 - u}\right), \quad K \geqslant 0. \qquad [19]$$

As in the shock probability experiment, discrimination probability in extinction is appropriately low only if bias heavily favors not responding. An interesting feature of Equation 19 is that discrimination probability is less than $b/(1 + b)$ when shock probability is greater than 0, and responding is only completely eliminated when the two shock schedules program shocks with probability 1. These values were not run by Herrnstein and Hineline, and the distinction would be a difficult one to observe in any case, since it would require distinguishing extremely low response rates from zero response rates.

While the results reported by the authors are qualitatively in line with the limiting behavior of p_D (Equations 17, 18, and 19), fits of mean interresponse and intershock time functions to their data are not feasible. Shock rates were not reported at all, and responding was reported in the form of *overall* rates, which included ineffective responding which occurred between the first postshock effective response and the next subsequent shock. Also, rate functions were averaged over different numbers of subjects, since not all subjects were run at all pairs of u_1, u_0 values. Thus a reliable assessment of the form of the response functions will have to await additional data. However, a rough fit of the data which are provided may be shown to require a low value of the bias parameter ($b < .1$), as in the preceding applications.

Avoidance Decrement

A puzzling phenomenon in the classical avoidance literature is the finding that under certain conditions rats show a marked decrement in avoidance performance over the course of extended acquisition training. The phenomenon has been found typically with wheel turning as the avoidance task, and is also related to sex, strain, and stock of the subjects. The effects of other variables such as shock intensity, warning signal duration, and intertrial interval are small (Anderson, 1969).

While the theory proposed here relates to asymptotic performance, the avoidance decrement phenomenon is similar to decreases in avoidance responding observed in the preceding applications when working at the avoidance task would still produce substantial improvement in intershock intervals. The dependence on genetic variables and especially on a particular response task implicates the bias parameter, b, which differentially weights responding and not responding. A descriptive approach to the decrement phenomenon is proposed below in which early performance is taken as an indicator of maximal avoidance efficiency, and later development of the decrement is viewed as the result of the biased discrimination process.

In Figure 8, a four-state chain for the classical avoidance paradigm is shown. After each shock (State S), discrimination trials are assumed to occur which favor working at the avoidance task with probability p_D. Shocks initiate an intertrial interval of r seconds followed by the warning signal (State W_S). Postshock and postresponse warning signals are represented by different states, since discrimination trials affect behavior only in the postshock warning signal. Warning signals are set to last for w seconds unless a response is made. If the discrimination trial after shock favors responding, subjects are assumed to time the postshock warning signal duration in the usual scalar fashion. Thus a response occurs in this signal with probability $p_D(1 - q)$. When the discrimination trial is unfavorable to responding, subjects simply wait out the signal for the next shock and discrimination trial. When a response does occur in the signal (State R), a different intertrial interval is programmed. Responses terminate the signal, but the signal timer, set at w seconds, is allowed to elapse before the "intertrial interval proper" begins. Thus the postresponse intertrial interval is r seconds plus $w - t_{WR}$, the amount of time "saved" by responding in the warning signal. After the postresponse intertrial interval elapses, a postresponse warning signal (W_R) occurs. Subjects are assumed to always attempt the avoidance task in this signal and thus responses occur with probability $1 - q$ and (scalar) latency $t_{WR} < w$. When a response is not made within w seconds, a shock is delivered and a new discrimination trial initiated. Frequently an escape contingency is included in shock which generates short escape latencies, so that the time between shock onset and the next warning signal (W_S) is actually variable and slightly longer than r. This additional time for postshock intertrial intervals is negligible for relatively long values of r, and is ignored in the following analysis.

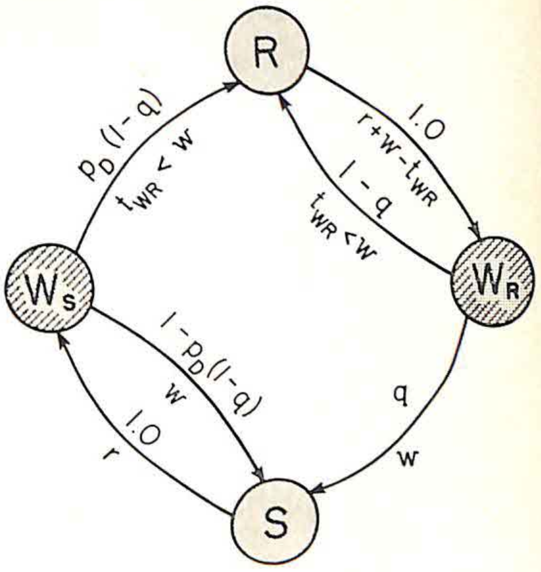


FIG. 8. State diagram for classical cued avoidance.

The discrimination analysis proceeds in the same way for the classical case as for the preceding free operant cases. Intershock intervals with no intervening responses are fixed at

$$x_0 = r + w. \qquad [20]$$

Intershock intervals including at least one response are obtained as follows: The time between onset of a postshock warning signal, W_S, and the next postresponse warning signal, W_R, is just $r + w$ seconds, since the postresponse intertrial interval compensates for the preceding response latency. Similarly when a response occurs in W_R, the time between warning signal onsets ($W_R \rightarrow R \rightarrow W_R$) is again $r + w$. Thus the intershock interval for responding may be written

$$x_1 = r + (r + w) + n(r + w) + w, \qquad [21]$$

where n is the (geometric) number of responses in W_R before the first shock.

In Appendix E1, x_1 is approximated in the usual way, and discrimination probability is found to take the simple form

$$p_D = \begin{cases} \exp\left[\dfrac{-q}{1-q}\left(\dfrac{1}{b} - 2\right)\right], & b < \frac{1}{2} \\ 1, & b \geqslant \frac{1}{2}. \end{cases} \qquad [22]$$

The striking feature of this formulation is that the temporal parameters r and w do not enter into discrimination probability at all. This results from the constant onset-to-onset time of the warning signals, which means that timing behavior in the signal is irrelevant to the form of the intershock interval distributions. The only timing parameter which counts is response probability, $1 - q$. According to Equation 22, discriminability of an improvement is perfect when $b \geqslant \frac{1}{2}$. This corresponds intuitively to the fact that even if only one response occurs between shocks, intershock time is doubled.

The avoidance decrement phenomenon consists of an early peak in avoidance performance followed by a subsequent decline. Assuming that the early peak represents the best that subjects can do, the maximum proportion of successful avoidance trials is an estimate of $1 - q$. A later decline in avoidance performance then depends on the development of $p_D < 1$ for an avoidance task for which $b < \frac{1}{2}$.

Asymptotic response probability is obtained in Appendix E2 as

$$\alpha_R = \frac{p_D(1-q)}{p_D(1-q) + q}. \qquad [23]$$

The limiting behavior of α_R with good and poor avoidance, and with bias, is described by the limits,

$$\alpha_R \rightarrow \begin{cases} 1 & \text{as } q \rightarrow 0, & [24a] \\ 0 & \text{as } q \rightarrow 1 \text{ or } b \rightarrow 0, & [24b] \\ 1 - q & \text{as } b \rightarrow \frac{1}{2}. & [24c] \end{cases}$$

Limit 24a means that a decrement should be difficult to discern if peak performance is very high, and Limit 24b means that the reverse is true if peak performance is low, or if bias heavily favors not responding. Limit 24c means that for any peak performance level, the decrement phenomenon should disappear as bias values approach $\frac{1}{2}$. A substantial decrement requires b considerably less than $\frac{1}{2}$, a range of values which is not unusual for the preceding applications. The decrement is *not* found for subjects from approximately the same genetic pool performing a shuttle task (Anderson, 1969; Nakamura & Anderson, 1964). This is consistent with the above formulation if the shuttle response is regarded as ethologically more "natural." The wheel-turning response appears, at least superficially, to have more in common with the lever response, and so bias values for these situations should be similar. Difficulties in training the discriminated lever response are familiar (D'Amato & Schiff, 1964; Meyer, Cho, & Wesemann, 1960), and a recent account (Bolles, 1970) provides extensive documentation for qualitative distinctions of this sort.

Assessment of Limits 24a, 24b, and 24c ideally requires experimental manipulation of bias or efficiency in the same subjects. A descriptive approach to this problem is effected in Figure 9 by plotting peak against asymptote from several different studies. The filled points are group averages for the wheel-turning task, and the open points are group averages for the shuttle task. The xs represent individual subjects with a lever-press task from an unpublished study by G. G. Neffinger conducted in this laboratory. Little or no decrement is observed with the shuttle task, as indicated by the open points along the diagonal; while substantial decrements may be seen in both the wheel and lever data. The smooth curves are plots of Equation 23 for the indicated values of b. Approximately 70% of the data fall between the curves for b values of .096 and .3, and all but one point within the range of .032 to .3. These values are comparable to those used to fit the free operant lever data in the preceding applications.

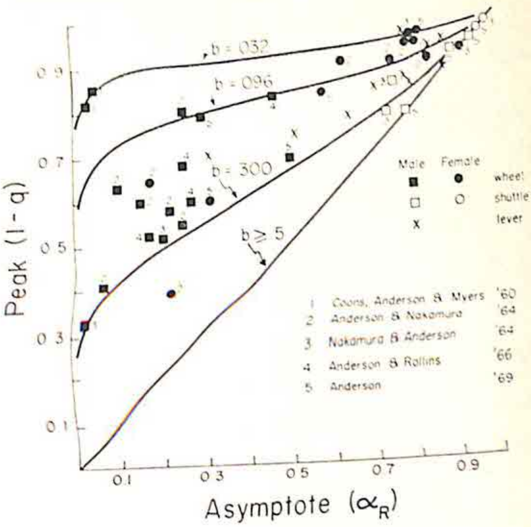


FIG. 9. Peak response probability $(1 - q)$ versus asymptotic response probability (α_R) for classical cued avoidance with a wheel-turning task (filled points) and a shuttle task (open points). (These data are group averages reported in the studies listed in the legend. Asymptotes were estimated by averaging response probabilities for the last two sessions run under the original training conditions irrespective of whether or not subjects received interpolated training under different conditions. The number of sessions run ranged from 12 to 40, and the number of trials per session ranged from 50 to 100. Subjects were rats from a similar genetic pool. The number of subjects per data point ranges from 6 to 16. The xs represent individual subjects with a discriminated lever-press response from an unpublished study by G. G. Neffinger. The temporal parameters were $r = 0$ seconds and $w = 20$ seconds. There were 360 trials per session and peaks were obtained from three-session averages (1,080 trials). Asymptotes were obtained from the final three sessions. The number of sessions ranged from 17 to 24. The theoretical curves describe the relation between $1 - q$ and α_R for the constant bias values indicated ("iso-bias contours").)

More substantial avoidance decrement for males than for females is often cited in these studies. In the context of Figure 9, most of the sex effect appears to result from generally higher peak performances for females (circles vs. squares). For instance, the original study reported a large decrement using male subjects (points in the upper left corner). In terms of the theory, these data represent an unusually strong bias in favor of not responding. However, an equally strong bias is required to describe the much smaller decrement observed in later studies for females with peak performances of about .95.

The theory is more congenial to single organism "own control" applications and to parametric manipulation at asymptote. Acquisition studies necessitate using groups which may vary in both relevant parameters, q and b. Thus the account is tentative. On the other hand, the discrimination formulation provides a provocative explanation for an otherwise puzzling phenomenon, and the analysis is appropriately sensitive to subject and response variables and insensitive to temporal parameter variables.

Shock Postponement

Recently, Hineline (1970) reported an important phenomenon in a cued "avoidancelike" paradigm. Responding by rats was acquired and maintained in a situation in which each response postponed but did not eliminate shock. The overall shock rate in each session was the same whether subjects responded or not (Experiment I). Maintained behavior in this situation was contrasted with a failure to obtain responding in a similar situation in which responses again postponed shock but also resulted in an *increase* of the overall shock rate (Experiment II). This pair of results taken together represents a considerable departure from the usual shock density finding. In the preceding applications, a considerable decrease in shock rate produced by responding was required for the maintenance of avoidance behavior. In Hine-

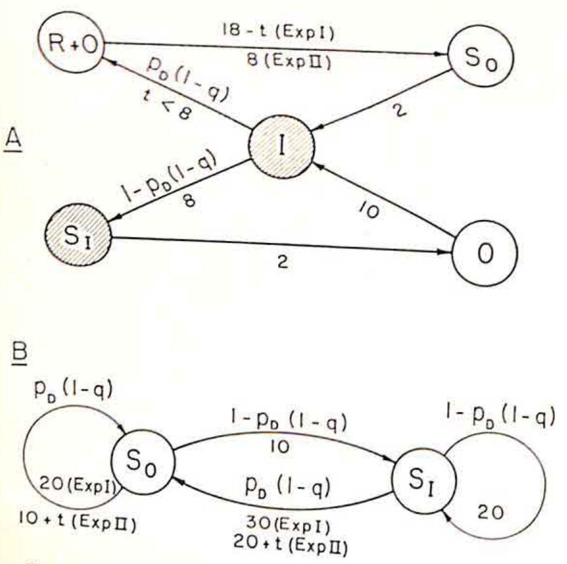


FIG. 10. State diagrams for the shock postponement experiments. The five-state chain (A) may be collapsed into the two-state chain (B) in which the states begin with shock delivery: either with the lever out (S_O) or in (S_I).

line's experiments, responding was maintained as long as it did not produce an *increase* in shock rate. The explanation proposed below hinges on the fact that responding altered *local* shock rates even when *overall* shock rates remained the same.

The manipulandum in these experiments was a retractable lever. Insertion of the lever into the experimental chamber initiated a trial, and a response produced immediate retraction of the lever. The situation was similar to a classical cued avoidance paradigm in that shock was delivered at the end of a fixed time after lever insertion when no response occurred. Shortly after this shock on nonresponse trials, the lever was withdrawn and remained out of the chamber for a fixed (intertrial interval) period. The situation differed from classical paradigms with respect to the consequences of responding. Responses terminated the trial by retracting the lever, but shock was always delivered in the following intertrial interval. Thus responses postponed, but did not "avoid," the shock which would otherwise have occurred during the trial.

In Figure 10A, a state diagram is presented for this situation. Insertion of the lever into the chamber is represented by State I (shaded). When no response occurred within eight seconds, a shock was delivered with the lever in (State S_I). Two seconds later the lever was retracted (State O), and remained out of the chamber for 10 seconds.[6] At the end of the 10-second out period, the lever was again inserted into the chamber and a new trial began. When a response occurred before eight seconds had elapsed ($t < 8$), the lever was retracted (State $R + O$), initiating a different intertrial interval. In Experiment I this intertrial interval was adjusted to "compensate" for response latency during the preceding trial, as in classical cued avoidance procedures (Figure 8). On response terminated trials, a shock was delivered 18 seconds after the trial had been initiated. This means that the time between a response and the subsequent shock delivered when the lever was out (State S_O) was variable ($18 - t$ seconds) and depended on response latency in the preceding trial. Two seconds after this intertrial interval shock, a new trial began with insertion of the lever into the chamber. The procedure in Experiment II differed from that in Experiment I in that the time between a response and the subsequent intertrial interval shock was held constant at eight seconds. This meant that responses resulted in an earlier delivery of the intertrial interval shock (S_O). The next trial began earlier also, with the result that overall shock frequency was increased substantially by responding. By contrast, in Experiment I the time between the beginning of one trial and the next was always 20 seconds, and just one shock was delivered in this period—either with the lever in or out.

Discrimination trials are assumed to occur after each shock, and the outcomes are reflected in response probability on the next trial. Thus subjects respond on any trial when the preceding discrimination favors responding (with probability p_D) and when the scalar timing process generates a response within the eight-second preshock interval (with probability $1 - q$). Otherwise [with probability $1 - p_D(1 - q)$] a response does not occur and shock is delivered after eight seconds with the lever in the chamber ($I \rightarrow S_I$).

---

[6] Responses could have occurred during the two seconds after shock delivery while the lever was still extended. However, Hineline reported that such responding was very rare, and so the $S_I \rightarrow R + O$ transition is ignored here.

The discrimination is based, as before, on a comparison of intershock intervals. However, in these experiments intershock intervals differ depending on whether shocks are delivered with the lever in or out of the chamber. This is clear in Figure 10B in which the five-state chain has been collapsed into a two-state chain by defining states as initiated by shock delivery, either with the lever out (S_O) or in (S_I). The four different transitions define four different intershock intervals. Two of these intervals contain an intervening response ($S_I \rightarrow S_O$ and $S_O \rightarrow S_O$), and for Experiment I these are greater than or equal to the two intershock intervals without a response ($S_O \rightarrow S_I$ and $S_I \rightarrow S_I$). For example, when a subject fails to respond for several trials in a row, shocks are delivered with the lever in ($S_I \rightarrow S_I$) at 20-second intervals. The first response results in an increase in the next intershock interval ($S_I \rightarrow S_O$). The additional shock-free time "earned" by this response is eventually "lost" by the first subsequent failure to respond which results in the shortest intershock interval of 10 seconds ($S_O \rightarrow S_I$). Thus in Experiment I there is a local differential in favor of intershock intervals containing responses. This differential might be experienced by subjects as quite substantial, depending on their performance and the way in which intervals are sampled for comparison. For example, a subject that rarely fails to respond on two successive trials experiences 10 seconds as the most frequent intershock interval for not responding. Both of the response intervals in Experiment I represent substantial improvement over this value.

The local differential in favor of responding in Experiment I is reduced in Experiment II. However, it is not clear that the reduction is sufficient to make intershock interval comparisons favor responding in Experiment I and not responding in Experiment II. Hineline reported asymptotic response probabilities in Experiment I ranging over approximately .75 to .95. Naive subjects in Experiment II showed some responding in the first few sessions, followed by a decline to near zero response probability. Two subjects run under both procedures showed maintained responding in Experiment I and an eventual drop to "negligible response rates" in Experiment II. Thus, while the data are not parametric, they demonstrate a clear qualitative distinction between the two procedures. The problem posed for the theory is whether a reasonable construction of the discrimination process is consistent with both outcomes.

In the preceding applications, samples of intershock intervals containing a response (x_1) were based on the distribution of this quantity produced by the scalar timing performance, since scalar timing was assumed to control responding after the first postshock response. Intershock intervals without a response (x_0) were independent of subjects' behavior. In the present paradigm, x_0 values are critically dependent on responding, since each response produces a current (long) x_1 and guarantees a future (short) x_0. The sampling assumption will be generalized for the shock postponement case by assuming that subjects sample x_1 and x_0 values from the distributions for both of these quantities produced by scalar timing alone. Setting p_D in Figure 10B equal to 1, the embedded discrete chain is stationary and the asymptotic relative frequency of each of the four transitions is just the asymptotic probability of observing the chain in the initial state times the probability of a one-step transition to the second state defining that interval. In Table 1, the four transitions are listed with their asymptotic probabilities. For example, an efficient timing performance with a high response probability $(1 - q)$ produces a high frequency of $S_O \rightarrow S_O(x_1)$ transitions, and a low frequency of $S_I \rightarrow S_I(x_0)$ transitions.

TABLE 1

TRANSITION TIMES AND PROBABILITIES FOR THE SHOCK POSTPONEMENT EXPERIMENTS

Transition $i \to j$	Time sample			Asymptotic probability $\alpha_i P(i \to j)$
	x	Experiment		
		I	II	
$S_I \to S_O$	x_1	30	$20 + t$	$q(1-q)$
$S_O \to S_O$	x_1	20	$10 + t$	$(1-q)^2$
$S_O \to S_I$	x_0	10	10	$q(1-q)$
$S_I \to S_I$	x_0	20	20	q^2

The response and nonresponse intershock intervals may be seen to reduce to Bernoulli samples [e.g., $P(x_0 = 20 | 10$ or $20) = q$], and discrimination probability is readily computed for x_1, x_0 pairs using the biased decision rule (Equation 2a). It is clear, however, that bias values comparable to those of the preceding applications with a nonretracting lever are not appropriate for Hineline's results. For example, values of b less than $\frac{1}{3}$ result in discrimination probabilities of zero in both experiments, since the largest differential is obtained when $x_1 = 30$ and $x_0 = 10$. Similarly, a high value of b under which responding is favored over not responding predicts maintained behavior in both experiments. Thus the problem for the theory becomes whether or not a range of reasonable values of b predicts maintained responding in Experiment I and negligible responding in Experiment II.

From Figure 10A, asymptotic response probability (α_R) is just the asymptotic probability of observing the discrete two-state chain in State S_O, because intertrial interval shocks occur if and only if a preceding response has occurred. Since the two-state chain (Figure 10B) is stationary,

$$\alpha_R \equiv \alpha_{S_O} = p_D(1 - q). \qquad [25]$$

In Figure 11, a discrimination of permissible bias values is effected by plotting possible asymptotes in Experiment I against possible asymptotes in Experiment II. The shaded area in the upper left corner represents a rough delineation of observed outcomes in Hineline's experiments. The stippled area above the equal-asymptote diagonal represents a rough delineation of expected outcomes from the two experiments, assuming an intermediate value of response latency ($t = 4$ seconds) and bias values either less than .5 or greater than .714. For bias values in the lower range, both asymptotes tend to be low, and for bias values in the upper range, both asymptotes tend to be high. Thus these ranges of b values are clearly disqualified by Hineline's results. The two curves passing through the shaded area represent expected outcomes for bias values between the upper and lower ranges. According to the theory, responding is high in Experiment *I* and low in Experiment II only if bias values for the retracting lever are in the range from .5 to approximately .7. This range is considerably higher than bias values in the preceding applications for the nonretracting lever in both cued and uncued situations. The discrepancy is sizable and deserves investigation.

Two observations support the view that the difference is real and represents different preferences for the two manipulanda. First, an earlier report (Hineline & Herrnstein, 1970) described responding in a classical paradigm (Figure 8) using the retracting lever, with asymptotic response probabilities that were extremely high (ap-

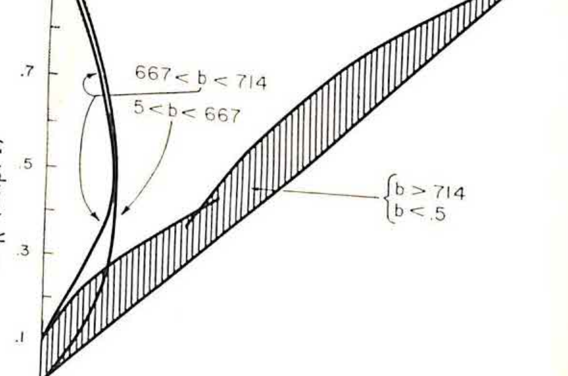


FIG. 11. Possible asymptotes (α_R) in the shock postponement experiments.

proximately .98). Such high values are very unlikely unless bias values exceed .5 (Equation 23). Moreover, the avoidance decrements observed by Neffinger with the nonretracting lever (Figure 9) were obtained using the same temporal parameters as in the Hineline and Herrnstein study. Second, I have attempted to train rats on a procedure similar to the shock postponement schedules using a nonretracting lever. Subjects have failed to acquire responding even with considerably larger x_1, x_0 differences than those used in Hineline's procedure for Experiment I. Valuable evidence on this view would be obtained by running subjects on these paradigms with and without lever retraction after each response.

The sampling assumptions are possibly more tenuous in this application than in the preceding cases, since values of discrimination probability less than 1.0 alter the relative frequency of x_1, x_0 values which subjects actually experience. The theory requires that the two distributions be maintained "in memory" at the frequency levels produced by the scalar performance. Alternative sampling assumptions might also produce the observed outcomes. For example, a subject that responded on every other trial would be expected to maintain this pattern in Experiment I if its b value was greater than $\frac{1}{3}$, and not respond at all in Experiment II if its b value was less than approximately .42.

At least one alternative sampling assumption is clearly ruled out by the data. If subjects are viewed as sampling shock rates over a substantial period of time (e.g., a session), and comparing rates when some responding occurs with rates when no responding occurs, then the procedure of Experiment I results in no difference. The point is an important one, since it means that any averaging in the sampling procedures must preserve the x_1, x_0 distinction between response and nonresponse intershock intervals. In the preceding application, this distinction was not forced, since a long time-sample with a high proportion of x_1 intervals might still be discriminated from a long time-sample of x_0 intervals. The shock postponement results, however, require a discrimination mechanism like that proposed here which is sensitive to local rather than overall shock density differences.

CONCLUDING DISCUSSION

The theoretical account proposed here holds that a variety of avoidance effects are attributable to the same source: confusability of intershock intervals produced by responding and not responding. The descriptive power of the account is difficult to fault as an overview. However, the details of the proposal may leave somewhat too much room for maneuver. Direct evidence for the assumptions underlying the discrimination process has not been obtained, and the data reported quite generally do not have the fine grain required to discriminate alternative models. For example, discrimination trials are assumed to occur directly after each shock. This placement of discrimination trials is convenient for the analysis of interresponse and intershock times, but it is not critical for a rough description of the data. It may be shown that assuming discrimination trials after responses as well as after shocks produces very similar predictions.

The discrimination analysis has been developed in the spirit of "ideal observer" detection models. A more realistic description might require that subjects *estimate* intershock intervals associated with responding and not responding. Such an estimation assumption would add variability to the real variance in the underlying distributions. However, the approximating techniques used to obtain discrimination probability invariably reduced the real variance in the intershock interval distributions, and the resulting discrimination probability functions were, if anything, too flat for an adequate description of the data. Thus a more realistic approach would also require averaging in the sampling procedures with the consequent addition of a subject parameter for sample size.

Bias

The logical extension of the bias analysis would be to manipulate bias independently of the parameters affecting timing. Bias values in avoidance situations with rats consistently favor not responding. However, the retracting lever and shuttle response tasks appear to be less "unnatural" than the standard lever and wheel-turning tasks. Discriminative feedback following the response may be implicated in this distinction. For example, it is common lore in free operant avoidance training that acquisition is facilitated by providing a distinctive click after every response. The retracting lever and shuttle tasks may maximize the distinctiveness of postresponse "safe signal" cues.

Parametric manipulation of bias is clearly desirable, but means to this end are not obvious. One possibility might be to increase the relative aversiveness of the response by increasing the effort required to execute it, or by punishing it with brief low-intensity shock. In the latter case, reversibility appears difficult to obtain (Powell & Morris, 1969). It is an empirical question, also, whether increasing the bias against responding might not put a motivation premium on spaced responding, thus sharpening the temporal discrimination of the preshock interval.

Another possibility, which has received some experimental attention, is to study choice in avoidance situations between two modes of responding, rather than between responding and not responding. Sidman and Boren (1957) and Sidman (1957) investigated a cued avoidance situation in which subjects appeared to choose the "more favorable" of two avoidance schedules at the cost of an associated shock. The typical situation was this: in the presence of a warning signal, subjects postponed shock by spacing their responses less than, say, w seconds apart. That is, each response recycled a timer set at w seconds. When subjects allowed w seconds to elapse without a response, a shock was delivered, the warning signal was terminated, and a different timer, set at v seconds, was initiated. Responses in the v-clock then recycled this timer until v seconds elapsed without a response. The warning signal then recurred and the w-timer was in effect until the next shock, and so on. The central finding for this situation was that when v was long relative to w, subjects responded more in the v-clock to postpone the warning signal than in the w-clock. Once the warning signal came on, subjects appeared to "take" the shock in order to return to the more favorable schedule in the absence of the signal.

This kind of choice represents a generalization of the discrimination analysis, since subjects appeared to be choosing between responding in one or the other clock rather than between responding and not responding. The discrimination process may be shown to handle the sort of preference described, if the decision is unbiased and represents a discrimination between shock density obtained when subjects are responding *only* in the *absence* of the signal, and shock density when subjects are responding *only* in the *presence* of the signal. Real evidence on the feasibility of this sort of a generalized discrimination process requires monitoring signal-onset and shock-onset probabilities in the two clocks, and these data are not reported. Also, the proposal for this situation assumes that the decision between responding of either sort and not responding at all is always made in favor of working at one or the other task. An alternative assumption might regard choices as being between three alternatives: not responding, or responding in either clock. Experimental arrangements of this sort are natural extensions of the two alternative response–nonresponse paradigms, and deserve detailed study in which direct evidence on changes in transition probabilities are obtained.

Two Factors versus Shock Density

It is possibly an important observation that a relatively broad range of phenomena in avoidance may be described theoretically without reference to a two-factor explanation. The reason, however, is at least in part substantive. Two-factor accounts are dynamic descriptions of an underlying mechanism presumed to maintain behavior. For example, Anger (1963) describes temporal discriminations as maintained by the difference between "dangerous" times close to shock delivery, and "safe" times further removed from shock delivery. The theoretical account proposed here is directed at understanding parameters of asymptotic behavior rather than understanding the way in which such behavior is acquired. The two approaches are certainly not mutually exclusive. The timing proposal, at least, seems congenial to a two-factor explanation, with responding supported by a drop in "conditioned aversiveness" or "fear" which is assumed to increase with proximity to shock-due time.

Theoretical accounts which challenge the primacy of two-factor explanations in favor of shock density as a motivating force (e.g., Herrnstein, 1969) often seem to overstate the opposition between the two views. The shock density position simply specifies an important variable controlling avoidance behavior, which is difficult to interpret in a two-factor context. However, a dynamic description of shock density effects is not thereby rendered irrelevant to an understanding of just how the shock density variable exerts control. For example, how is conditioned aversiveness distributed over the intervals between shocks? Are long intershock intervals or low shock rates reinforcing due to less aversiveness per unit time?—that is, due to "aversiveness density"? And how is improvement translated into behavior? The discrimination process proposed here provides answers to questions of this sort which are probably overly simplistic. However, the theory is sufficiently detailed that the assumptions have testable consequences. The account thus stands a chance of being demonstrably wrong, a virtue which is less easily claimed for more qualitative approaches to shock density effects.

A two-factor interpretation of scalar timing would also require substantial elaboration. While it is feasible to view responding as supported by a consequent drop in conditioned aversiveness, parsimony is not served by a one-to-one correspondence of this sort, and it is not at all clear how the scalar property follows. One approach might attempt to distinguish between the *perception* and the *experience* of danger. On this view, the perceptual component of timing generates subjective estimates of the proximity of the aversive event, and different subjective proximities are associated with different levels of the experience of fear. The Weber-like scalar property may then be restricted to the perceptual rather than the motivational component of timing. While this approach seems intriguing and even reasonable, sharp experimental referents are not immediately apparent.

REFERENCES

Anderson, N. H. Variation of CS–US interval in long-term avoidance conditioning in the rat with wheel turn and with shuttle tasks. *Journal of Comparative and Physiological Psychology*, 1969, **68**, 100–106.

Anderson, N. H., & Nakamura, C. Y. Avoidance decrement in avoidance conditioning. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 196–204.

Anderson, N. H., & Rollins, H. A. Two failures to prevent avoidance decrement. *Psychological Reports*, 1966, **19**, 71–78.

Anger, D. The role of temporal discriminations in the reinforcement of Sidman avoidance behavior. *Journal of the Experimental Analysis of Behavior*, 1963, **6** (Suppl.), 477–506.

Bolles, R. C. Species-specific defense reactions and avoidance learning. *Psychological Review*, 1970, **77**, 32–48.

Boren, J. J., & Sidman, M. Maintenance of avoidance behavior with intermittent shocks. *Canadian Journal of Psychology*, 1957, **11**, 185–192.

Coons, E. E., Anderson, N. H., & Meyers, A. K. Disappearance of avoidance responding during continued training. *Journal of Comparative and Physiological Psychology*, 1960, **53**, 290–292.

D'Amato, M. R., & Schiff, D. Long-term discriminated avoidance performance in the rat. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 123–126.

Feller, W. *An introduction to probability theory and its applications*. Vol. 2. New York: Wiley, 1966.

Gibbon, J. Scalar timing and semi-Markov chains in free-operant avoidance. *Journal of Mathematical Psychology*, 1971, **8**, 109–138.

Herrnstein, R. J. Method and theory in the study of avoidance. *Psychological Review*, 1969, **76**, 49–69.

Herrnstein, R. J., & Hineline, P. N. Negative reinforcement as shock-frequency reduction. *Journal of the Experimental Analysis of Behavior*, 1966, **9**, 421–430.

Hineline, P. N. Negative reinforcement without shock reduction. *Journal of the Experimental Analysis of Behavior*, 1970, **14**, 259–268.

HINELINE, P. N., & HERRNSTEIN, R. J. Timing in free-operant and discrete-trial avoidance. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 113–126.

McGILL, W. J., & GIBBON J. The General-Gamma distribution and reaction times. *Journal of Mathematical Psychology*, 1965, **2**, 1–18.

MEYER, D. R., CHO, C., & WESEMANN, A. F. On problems of conditioning discriminated lever-press avoidance responses. *Psychological Review*, 1960, **67**, 224–228.

NAKAMURA, C. Y., & ANDERSON, N. H. Avoidance conditioning in wheel box and shuttle box. *Psychological Reports*, 1964, **14**, 327–334.

POWELL, R. W., & MORRIS, G. Continuous punishment of free-operant avoidance in the rat. *Journal of the Experimental Analysis of Behavior*, 1969, **12**, 149–157.

SHEFFIELD, F. D. Avoidance training and the contiguity principle. *Journal of Comparative and Physiological Psychology*, 1948, **41**, 165–177.

SHNIDMAN, S. R. Extinction of Sidman avoidance behavior. *Journal of the Experimental Analysis of Behavior*, 1968, **11**, 153–156.

SIDMAN, M. Two temporal parameters of the maintenance of avoidance behavior by the white rat. *Journal of Comparative and Physiological Psychology*, 1953, **46**, 253–261.

SIDMAN, M. The temporal distribution of avoidance responses. *Journal of Comparative and Physiological Psychology*, 1954, **47**, 399–402.

SIDMAN, M. Conditioned reinforcing and aversive stimuli in an avoidance situation. *Transactions of the New York Academy of Science*, 1957, **19**, 534–544.

SIDMAN, M. Reduction of shock frequency as reinforcement for avoidance behavior. *Journal of the Experimental Analysis of Behavior*, 1962, **5**, 247–257.

SIDMAN, M. Avoidance behavior. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application.* New York: Appleton-Century-Crofts, 1966.

SIDMAN, M., & BOREN, J. J. A comparison of two types of warning stimulus in an avoidance situation. *Journal of Comparative and Physiological Psychology*, 1957, **50**, 282–287.

WERTHEIM, G. A. Some sequential aspects of IRT's emitted during Sidman-avoidance behavior in the white rat. *Journal of the Experimental Analysis of Behavior*, 1965, **8**, 9–15.

APPENDIXES

APPENDIX A

N-STATE SEMI-MARKOV CHAINS

Let $\mathbf{c} = \{c_{ij}\}$, $i, j = 1, 2, \cdots, N$ be the transition matrix for the embedded discrete Markov chain. (See Gibbon, 1971, for a more extensive treatment of the relevant theory in Markov and semi-Markov processes.) Assuming that $\mathbf{c}$ is regular, the discrete chain has a limiting probability (fixed) vector, $\boldsymbol{\alpha} = \{\alpha_j\}$, with

$$\lim_{n\to\infty} \mathbf{c}^n = \mathbf{1}\boldsymbol{\alpha},$$

where $\mathbf{1}$ is the column vector with all elements equal unity. $\boldsymbol{\alpha}$ is the unique probability vector with the property that

$$\boldsymbol{\alpha}\mathbf{c} = \boldsymbol{\alpha}. \qquad [\text{A1}]$$

For example, the two-state chain in Figure 2B has

$$\mathbf{c} = \begin{array}{c} \\ R \\ S \end{array} \!\!\begin{array}{cc} R & S \\ \left(\begin{array}{c} 1-q \\ p_D(1-q) \end{array}\right. & \left.\begin{array}{c} q \\ 1-p_D(1-q) \end{array}\right) \end{array}, \qquad [\text{A2a}]$$

and the solution of Equation A1 for the two-state chain is well known to be $\alpha_j = c_{ij}/(c_{ij}+c_{ji})$, $i \neq j = 1, 2$, so that

$$\alpha_R = \frac{p_D(1-q)}{p_D(1-q)+q}, \quad \alpha_S = 1-\alpha_R. \qquad [\text{A2b}]$$

Let μ_{ij} be the mean one-step transition time from State i to State j, and define the mean unconditional exit time from State i as

$$\bar{\mu}_i = \sum_{j=1}^{N} c_{ij}\mu_{ij}. \qquad [\text{A3}]$$

Then two basic results for the semi-Markov chain are:

$$\pi_j = \frac{\alpha_j\bar{\mu}_j}{\sum_{k=1}^{N} \alpha_k\bar{\mu}_k}, \qquad [\text{A4}]$$

and

$$m_{jj} = \bar{\mu}_j/\pi_j, \qquad [\text{A5}]$$

where π_j is the asymptotic state-occupancy probability for State j, and m_{jj} is the mean (unconditional) recurrence time for State j. Equations A4 and A5 are the analogues in the semi-Markov process of α_j and α_j^{-1} for the discrete chain.

For the two-state chain, the scalar property (Equation 1a) applied to the mean latency in both clocks, gives

$$\bar{\mu}_R = r[1 - (1 - q)(1 - \mu)],$$
$$\bar{\mu}_S = s[1 - p_D(1 - q)(1 - \mu)]$$

for the mean exit time from each state. Mean recurrence times may then be written

$$m_{RR} = r[\mu + q(1 - \mu)] + sq\left[\frac{1}{p_D(1 - q)} + 1 - \mu\right] \qquad [\text{A6a}]$$

$$m_{SS} = \left[\frac{p_D(1 - q)}{q}\right] m_{RR}. \qquad [\text{A6b}]$$

APPENDIX B

p_D AS A FUNCTION OF r, s, IN UNCUED AVOIDANCE

The time between shocks including at least one response is given by Equation 3 as $x = t_{SR} + \tau + r$. Using the scalar property for means (Equation 1a), the expected value of τ (Equation 4), according to Equation 5, is

$$E(\tau) = r\mu(1 - q)/q \equiv \lambda^{-1}. \qquad [\text{B1}]$$

τ may be approximated by an exponential with a rate parameter, λ, equal to the inverse of Equation B1. Since t_{SR} contributes little variance to x relative to variance in τ, it is approximated by its mean, $s\mu$, giving

$$x \simeq s\mu + \tau + r. \qquad [\text{B2}]$$

The biased decision rule, Equation 2, may be recast in terms of the approximation, Equation B2, as $\tau > s[(1/b) - \mu] - r$. Discrimination probability, p_D, is then the tail area of the exponential with rate parameter λ. Some algebraic rearrangement shows p_D to be the function of the ratio of r and s given by Equation 6.

Under the scalar assumption, mean transition times are given by

$$E\{t_{ij}\} = \begin{array}{c|ccc} & R & S^+ & S^- \\ \hline R & r\mu & r & r \\ S^+ & s\mu & s & s \\ S^- & & s & s \end{array}.$$

APPENDIX C

p_D (1) AND m_{jj} (2) AS FUNCTIONS OF SHOCK PROBABILITY

1. It is convenient to restate the decision rule as $x_1 > x_0/b$ and approximate the right side as

$$x_0/b \simeq s/b + \tau_0, \qquad [\text{C1}]$$

where τ_0 is exponential with rate parameter,

$$\lambda_0 = E(\tau_0)^{-1} = bu/[s(1 - u)].$$

Intershock intervals when responding (Equation 10) are approximated analogously to the preceding case as,

$$x_1 \simeq s\mu + \tau_1 + r + \tau_2, \qquad [\text{C2}]$$

where τ_1, τ_2 are exponential with rate parameters,

$$\lambda_1 = q/[r\mu(1 - q)],$$

and

$$\lambda_2 = u/[s(1 - u)].$$

Using these approximations, the decision rule may be rewritten as

$$\tau_1 + \tau_2 > \tau_0 + K \qquad [\text{C3}]$$

where $K = s(1/b - \mu) - r$. The sum on the left of the inequality, C3, is distributed as a twofold general gamma variate (McGill & Gibbon, 1965), and the right side variate is exponential. Discrimination probability is thus the area to the right of K in the difference distribution of the variates on either side of the inequality. A straightforward application of the convolution integral shows p_D to be given by Equation 11.

2. Mean interresponse and intershock times are found from the general recurrence time formulation (Equation A5) by combining the shock (S^+) and discriminate (D) states as before. This results in the following transition matrix for the embedded discrete chain:

$$\mathbf{c} = \begin{array}{c|ccc} & R & S^+ & S^- \\ \hline R & 1 - q & qu & q(1 - u) \\ S^+ & p_D(1 - q) & u[1 - p_D(1 - q)] & [1 - u][1 - p_D(1 - q)] \\ S^- & 0 & u & 1 - u \end{array}.$$

The $\mathbf{c}$ matrix is regular, since $\mathbf{c}^2$ is positive. The limiting state probabilities may therefore be found by solving Equation A1. The result may be written

$$\alpha_j = \frac{K_j}{\sum_i K_i}, \quad i, j = R, S^+, S^-, \qquad [\text{C4}]$$

where $K_j = up_D(1 - q)$, qu, $q(1 - u)$, for $j = R, S^+, S^-$, respectively.

In the Boren and Sidman (1957) experiment, r was equal to s. Obtaining the $\bar{\mu}_j$ according to Equation A3 with $r = s$, and using Solution C4, the mean interresponse and intershock time formulations may be reduced to the simple forms of Equations 13a and 13b.

APPENDIX D

p_D FOR DIFFERENTIAL SHOCK PROBABILITY

As before, the decision rule may be recast and divided by $r = s$ to give

$$x_1/s > x_0/sb, \qquad [D1]$$

Approximating the geometric random variables in x_0, x_1 by exponentials, the right and left sides of Inequality D1 become (from Equations 14 and 15),

$$x_0/sb \simeq \tau_0 + 1/b$$

and

$$x_1/s \simeq \mu + \tau_1 + 1,$$

where τ_0, τ_1 are exponentials with rate parameters, $\lambda_0 = bu_0/(1 - u_0)$, and $\lambda_1 = u_1/(1 - u_1)$. Inequality D1 may then be written,

$$\tau_1 > \tau_0 + K, \qquad [D2]$$

where $K = 1/b - \mu - 1$. Decision Rule D2 is a special case of C3, for $\tau_2 = 0$. Thus p_D may be obtained from Equation 11 by letting $\lambda_2 \rightarrow \infty$. The result is Equation 16.

APPENDIX E

p_D (1) AND α_R (2) FOR CLASSICAL CUED AVOIDANCE

1. The decision rule may be recast for this case as,

$$x_1/(r + w) > x_0/b(r + w).$$

From Equation 20 the right side is just $1/b$, while the left side may be approximated as,

$$x_1/(r + w) \simeq 2 + \tau,$$

where τ is exponential with rate constant, $\lambda = q/(1 - q)$. Discrimination probability is then the area to the right of $1/b - 2$, as given by Equation 22.

2. The four-state chain in Figure 9 is not Markov since transition times are correlated, and in any case, the embedded discrete chain is periodic. However, the response and shock states may be combined with the following warning signal states, W_R and W_S. This results in a discrete chain in which all transition times are $r + w$ seconds and transition probabilities are

$$\mathbf{C} = \begin{array}{c} \\ W_R \\ W_S \end{array} \begin{array}{cc} W_R & W_S \\ \left[\begin{array}{c} 1 - q \\ p_D(1 - q) \end{array}\right. & \left.\begin{array}{c} q \\ 1 - p_D(1 - q) \end{array}\right] \end{array},$$

as in the two-state uncued avoidance chain (Equation A2a). The data reported do not discriminate between trials following shocks and trials following avoidance responses, so that asymptotic response probability is a weighted sum of response probability in each kind of trial, with the weights being, α_j, $j = W_R$, W_S, the long-range relative frequencies of the two kinds of trial. But response probability in W_R is the probability of the $W_R \rightarrow W_R$ transition, and in W_S it is the probability of the $W_S \rightarrow W_R$ transition. Then by the vector property, Equation A1,

$$P(W_R \rightarrow W_R)\alpha_{W_R} + P(W_S \rightarrow W_R)\alpha_{W_S} = \alpha_{W_R}.$$

Asymptotic avoidance is just the asymptotic probability of observing the two-state chain in State W_R. From the uncued avoidance formulation, Equation A2b, setting $\alpha_{W_R} \equiv \alpha_R$ results in Equation 23.

(Received June 4, 1971)

Psychological Review
1972, Vol. 79, No. 1, 93-96

THEORETICAL NOTE

NEGATIVE AS WELL AS POSITIVE SYNAPTIC CHANGES MAY STORE MEMORY[1]

MARK R. ROSENZWEIG[2]
University of California, Berkeley

KJELD MØLLGAARD
University of Copenhagen, Denmark

MARIAN C. DIAMOND AND EDWARD L. BENNETT
University of California, Berkeley

Hypotheses concerning synaptic changes that might store memory have been appearing ever since E. Tanzi proposed the first one in 1893. Most of them have been cast in terms of positive changes with learning, that is, either growth of new synaptic junctions or changes in size or effectiveness of existing ones. In the last decade, a few theorists have proposed that negative changes—decreases in the number or effectiveness of contacts—could store memory. Little data have been available to test either hypothesis. We propose that structural synaptic changes underlying long-term memory need not be of a single form but may include increase in number of synapses, increase in size, decrease in number, and decrease in size. Recent electron microscopic data are reviewed to show that each of these four types of change has been observed as a consequence of differential experience.

Shortly after the neuron doctrine was enunciated, the psychiatrist Eugenio Tanzi (1893) surveyed recent advances in neurohistology and hypothesized that learning produces anatomical changes in interneuronal contacts (which were to be named "synapses" by Foster & Sherrington, 1897), and that these changes serve as a mechanism of memory. Tanzi was confident that investigators would soon be able to test by direct inspection the changes that he hypothesized to occur with age and training. Now, almost 80 years later, researchers are just beginning to be able to perform such tests (Bloom, 1970; Kandel & Spencer, 1968). While some have become pessimistic ("No major breakthrough in the synaptic connectivity hypothesis of memory has been achieved, despite extensive work during the past 10 years . . . [Baumgarten, 1970, p. 264]"), others decry that more has not been done to link various cerebral changes in learning to synaptic mechanisms (Anonymous, 1971).

Most theorists during the intervening period, whether or not they knew of Tanzi, have supposed with him that learning involves either the formation of new interneuronal connections or the strengthening of existing ones; few have thought of learning in terms of elimination or weakening of synapses. Foster and Sherrington (1897, Pt. III), also related this structure to learning:

> Shut off from all opportunity of reproducing itself and adding to its number by mitosis or otherwise, the nerve cell directs its pent-up energy towards amplifying its connections with its fellows, in response to the events which stir it up. Hence, it is capable of an education unknown to other tissues [p. 1117].

Such concepts were taken up by a number of theorists during the present century and were brought to prominence in the late 1940s by Konorski (1948) and Hebb (1949). Hebb assumed explicitly that "the changed facilitation that constitutes learning" is based on "the growth of synaptic knobs [p. 65]." Formulations of some authors have not made clear whether growth in the nervous system with learning involves formation of new synapses or changes in size or other dimensions of existing synapses. Other workers have discriminated between these possibilities, as when Eccles (1965) stated his belief that learning and memory storage involve "growth just of bigger and better synapses that are already there, not growth of new connections."

The elimination or weakening of synaptic connections as a mechanism of learning and memory storage has been considered much less frequently than have positive changes. Concerning embryological development of the nerv-

---

[1] This research was supported by National Science Foundation Grant GB-8011 and also received support from the United States Atomic Energy Commission.

[2] Requests for reprints should be sent to Mark R. Rosenzweig, Department of Psychology, Tolman Hall, University of California, Berkeley, California 94720.

ous system, Ramón y Cajal (cited by Jacobson, 1970) did propose that there is a selection among somewhat random original connections "due to atrophy of certain collaterals and the progressive disappearance of disconnected or useless neurones [p. 125]." A large decrease in spinal neurons during embryology has been demonstrated in later work (Jacobson, 1970). While it has been suggested that "synaptic atrophy" (Konorski, 1948) or "synaptic decay" (Hebb, 1949) may occur with disuse even in the adult, this was proposed as a correlate of forgetting, the converse of acquisition of memory. Even Harlow (1959; Harlow & Hicks, 1957), who has suggested that learning is accomplished largely through the reduction of errors, conceived of this as occurring through the establishment of inhibitory processes rather than through elimination of excitatory ones (see Harlow & Hicks, 1957, pp. 108–109). Perhaps the prevalence of notions of positive synaptic changes with learning stems from the concept of "consolidation of memory traces" formulated in 1900 by Müller and Pilzecker. The first to hypothesize negative changes to account for learning may have been Thesleff (1962). He suggested that, in analogy to facts established for muscle cells, the chemoreceptor area of a neuron might become restricted to sites activated frequently, and that other synaptic connections would thereby be rendered inoperative. Young (1964, p. 285) suggested that learning occurs through the reduction of unused redundant connections, but he did not elaborate this hypothesis. Recently, Dawkins (1971) proposed that selective death of neurons may represent a mechanism of memory storage. These suggestions have remained speculative because, until very recently, there were no data against which to test them.

We hypothesize that structural synaptic changes underlying learning and memory storage need not be of a single form, but rather may take one or more of these four forms: (*a*) increase in number of synapses, (*b*) increase in size of contact areas, (*c*) decrease in number, and (*d*) decrease in size of contact. It would be expected that with any particular instance of learning, such changes would be limited to particular brain regions and to particular types of contacts; in fact, different changes may occur simultaneously in different sets of synapses in the same brain.

A few examples from recent studies will serve to demonstrate that changes of each of these sorts are beginning to be reported and that more than one type of change may occur within the same region of the brain. Cragg (1967, 1968, 1969) has made several reports on measures of axon terminals in rats kept permanently in the dark or removed for various periods of light exposure. In 1967, he reported findings suggestive of all four types of synaptic changes: When dark-reared rats were exposed to daylight, axon terminals in the upper half of the occipital cortex became larger and less numerous, whereas in the lower half they became smaller and more numerous. His 1969 report states that axon terminals in the lateral geniculate nucleus are 15% greater in diameter in dark-raised rats, but 34% less in density per unit of tissue volume. When he measured length of synaptic contacts rather than axon terminals, the contacts "did not show any regular difference between the light- and dark-exposed rats of successive pairs [p. 61]." Fifková (1970) measured axodendritic synapses in Layers II–IV of occipital cortex of rats after six weeks of unilateral lid suture; synapses in the hemisphere contralateral to the sutured eye were 20% fewer in number but 7.5% larger in cross-section than in the control hemisphere.

We have been using quite a different procedure, including enriched or impoverished experience, which induces a variety of chemical and anatomical changes in the rodent brain (Rosenzweig, Bennett, & Diamond, in press), and we have recently found that this produces clear synaptic changes (Møllgaard, Diamond, Bennett, Rosenzweig, & Lindner, 1971). In our procedure, the main factor inducing cerebral effects is experience with a variety of stimulus objects in the enriched condition; the impoverished-experience animals are isolated but are not deprived of visual stimulation. The experimental conditions present opportunities for differential experience over several weeks. Control experiments have shown that the cerebral effects previously measured are not attributable to stress, locomotion, altered rates of maturation, endocrine involvement, handling, or to any brief irrelevant stimulation. Visual experience is not required to produce these effects, since they occur with groups run completely in the dark and also in blinded rats (Rosenzweig, Bennett, Diamond, Wu, Slagle, & Saffran, 1969). The effects are not closely age dependent, since some at least have been induced in 300-day-old rats (Riege, 1971).

Significant differences in size and number of synaptic junctions were found between littermate rats assigned at weaning (25 days of age) to enriched or impoverished environments and kept there for 30 days. The synapses measured were asymmetrical axodendritic junctions in the neuropil of Layer III of the occipital cortex. Rats from the enriched condition showed, in comparison to littermates from the impoverished condition, synapses that averaged 48% greater in length but that were only 67% as numerous. The enriched rats had more large synapses as well as fewer small synapses than did impoverished rats, so the enriched condition size distribution could not have been derived simply by loss of small synapses from the impoverished condition distribution. The total area of synapses in the enriched group, taking both size and number of contacts into account, was 40% greater than in the impoverished group. It should be noted that the inverse effects between size and number of synapses in this study is similar to the relation between synaptic size and number reported by Cragg and Fifková.

In another study,[3] evidence of a positive enriched-impoverished difference in a number of dendritic spines (and thus presumably of synaptic contacts) on pyramidal cells in the occipital cortex of rats was obtained. Counts of spines were made on enriched-impoverished littermate pairs from four experiments with behavioral conditions similar to those of the preceding paragraph. On basal dendrites there was a 9.7% enriched-impoverished difference ($p < .01$) in the number of spines per unit of length along the dendrite. The counts of spines by light microscopy on cells stained by the Golgi-Cox method cannot be assigned to cortical layers for direct comparison with the preceding electron microscopic results. It is possible that electron microscopic counts in other cortical layers will reveal an increase in the number of junctions, or the relation between electron microscopic synaptic counts and Golgi-Cox spine counts may be more complex; only further research can answer this question. When enriched-experience rats are compared with impoverished-experience littermates, they appear to have a lesser number but greater size in one population of synapses and to have a greater number in another population. Experimental evidence from several laboratories is thus consistent with the old hypotheses that stimulation and learning may induce increases in synaptic size and number, and it also suggests that negative as well as positive changes occur.

Our measures of synapses and of dendritic spines were made in occipital cortex because that is where we have found the largest changes in other anatomical and biochemical measures after several weeks of differential experience (Rosenzweig et al., 1969). Experiments with brief training by other investigators have shown chemical changes located especially in the hypothalamus (Beach, Emmens, Kimble, & Lickey, 1969; Kahan, Kriginan, Wilson, & Glassman, 1970). Some groups (e.g., Halas, Beardsley, & Sandlie, 1970) are recording the activity of single neurons at various levels of the brain during conditioning. Results indicate that the earliest changes during training occur in the reticular formation and in basal reward areas, with changes occurring eventually in the cerebral cortex. Thus, the chemical and electrophysiological observations both suggest a shift of locus from short-term effects involving motivational or reinforcing areas in basal regions of the brain to long-term storage, especially in the cerebral cortex. The chemical and anatomical changes that we find at the cortex may well represent the cumulation of innumerable small effects of learning during the prolonged differential experience.

It should be clear that we are not hypothesizing that the four types of synaptic changes are the only mechanisms for storing memory. In particular, some memories are formed very rapidly, and at least the initial type of storage is unlikely to be a structural change of synapses (although a substantial increase in the proportion of large axon terminals was found after 65 minutes of stimulation in the spinal cord of cats; Illis, 1969). Long-term memory storage, however, could well involve the positive and negative synaptic changes described here.[4]

---

[3] A. Globus, M. R. Rosenzweig, E. L. Bennett, and M. C. Diamond. Effects of differential environment on dendritic spine counts. In preparation.

[4] After completion of this paper we found somewhat related hypotheses in Albus (1971). Albus holds that in order for the learning process to be stable in cerebellar cells, and possibly elsewhere in the nervous system, "pattern storage must be accomplished principally by weakening synaptic weights rather than by strengthening them [p. 25]." He notes that "The argument synaptic weights are weakened by learning rather than strengthened is counterintuitive and contrary to most, if not all, theories of synaptic learning that have appeared in the literature [p. 50]," and he therefore examines it in detail. Later, he holds that synaptic facilitation is a mechanism of learning in stellate cells and elsewhere (p. 52). It should be noted that the physiological or anatomical mechanisms by which "synaptic weights" are expressed or altered are never defined. Nevertheless, it is interesting that Albus, arguing from quite different material, suggests as we do that both negative nad positive synaptic changes may store memory.

REFERENCES

Albus, J. S. A theory of cerebellar function. *Mathematical Biosciences*, 1971, **10**, 25–61.

Anonymous. Why do brain cells synthesize proteins? *Nature New Biology*, 1971, **230**, 100.

Baumgarten, R. J. von. Plasticity in the nervous system at the unitary level. In F. O. Schmitt (Ed.), *The neurosciences. Second study program.* New York: Rockefeller University Press, 1970.

Beach, G., Emmens, M., Kimble, D., & Lickey, M. Autoradiographic demonstration of biochemical changes in the limbic system during avoidance training. *Proceedings of the National Academy of Science*, 1969, **62**, 692–696.

Bloom, F. E. Correlating structure and function of synaptic ultrastructure. In F. O. Schmitt (Ed.), *The neurosciences. Second study program.* New York: Rockefeller University Press, 1970.

Cragg, B. G. Changes in visual cortex on first exposure of rats to light: Effect on synaptic dimensions. *Nature*, 1967, **215**, 251–253.

Cragg, B. G. Are there structural alterations in synapses related to functioniong? *Proceedings of the Royal Society, Series B*, 1968, **171**, 319–323.

Cragg, B. G. The effects of vision and dark-rearing on the size and density of synapses in the lateral geniculate nucleus measured by electron microscopy. *Brain Research*, 1969, **13**, 53–67.

Dawkins, R. Selective neurone death as a possible memory mechanism. *Nature*, 1971, **229**, 118–119.

Eccles, J. C. Possible ways in which synaptic mechanisms participate in learning, remembering and forgetting. In D. P. Kimble (Ed.), *The anatomy of memory.* Palo Alto, Calif.: Science & Behavior Books, 1965.

Fifková, E. The effect of monocular deprivation of the synaptic contacts of the visual cortex. *Journal of Neurobiology*, 1970, **1**, 285–295.

Foster, M., & Sherrington, C. S. *A text book of physiology.* Part III. *The central nervous system.* New York: Macmillan, 1897.

Halas, E. S., Beardsley, J. V., & Sandlie, M. E. Conditioned neuronal responses at various levels in conditioning paradigms. *Electroencephalography and Clinical Neurophysiology*, 1970, **28**, 468–477.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science.* Study 1. *Conceptual and systematic.* Vol. 2. *General systematic formulations, learning, and special processes.* New York: McGraw-Hill, 1959.

Harlow, H., & Hicks, L. H. Discrimination learning theory: Uniprocess vs. duoprocess. *Psychological Review*, 1957, **64**, 104–109.

Hebb, D. O. *The organization of behavior.* New York: Wiley, 1949.

Illis, L. S. Enlargement of spinal cord synapses after repetitive stimulation of a single posterior root. *Nature*, 1969, **223**, 76–77.

Jacobson, M. Development, specification, and diversification of neuronal connections. In F. O. Schmitt (Ed.), *The neurosciences. Second study program.* New York: Rockefeller University Press, 1970.

Kahan, B. E., Kriginan, M. R., Wilson, J. E., & Glassman, E. Brain function and macromolecules. VI. Autoradiographic analysis of the effect of a brief training experience on the incorporation of uridine into mouse brain. *Proceedings of the National Academy of Science*, 1970, **65**, 300–304.

Kandel, E. R., & Spencer, W. A. Cellular neurophysiological approaches in the study of learning. *Physiological Review*, 1968, **48**, 65–134.

Konorski, J. *Conditioned reflexes and neuron organization.* Cambridge: Cambridge University Press, 1948.

Møllgaard, K., Diamond, M. C., Bennett, E. L., Rosenzweig, M. R., & Lindner, B. Quantitative synaptic changes with differential experience in rat brain. *International Journal of Neuroscience*, 1971, **2**, 113–128.

Müller, G. E., & Pilzecker, A. Experiementelle Beiträge zur Lehre vom Gedächtnis. *Zeitschrift für Psychologie*, 1900, Ergbd. I, 1–288.

Riege, W. H. Environmental influences on brain and behavior of old rats. *Developmental Psychobiology*, 1971, **4**, 157–167.

Rosenzweig, M. R., Bennett, E. L., & Diamond, M. C. Chemical and anatomical plasticity of brain: Replications and extensions, 1970. In J. Gaito (Ed.), *Macromolecules and behavior.* (2nd ed.) New York: Appleton-Century-Crofts, in press.

Rosenzweig, M. R., Bennett, E. L., Diamond, M. C., Wu, S-Y., Slagle, R. W., & Saffran, E. Influences of environmental complexity and visual stimulation on development of occipital cortex in rat. *Brain Research*, 1969, **14**, 427–445.

Tanzi, E. I fatti e le induzioni nell'odierna istologia del sistema nervoso. *Rivista Sperimentale di Freniatria e di Medicina Legale*, 1893, **19**, 419–472.

Thesleff, S. A neurophysiological speculation concerning learning. *Perspectives in biology and medicine*, 1962, **5**, 293–295.

Young, J. Z. *A model of the brain.* Oxford: Clarendon Press, 1964.

(Received June 18, 1971)

Page 1

U. S. POSTAL SERVICE

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

(Act of August 12, 1970: Section 3685, Title 39, United States Code)

SEE INSTRUCTIONS ON PAGE 2 (REVERSE)

1. TITLE OF PUBLICATION
PSYCHOLOGICAL REVIEW

2. DATE OF FILING
9/28/71

3. FREQUENCY OF ISSUE
BI-MONTHLY

4. LOCATION OF KNOWN OFFICE OF PUBLICATION *(Street, city, county, state, ZIP code) (Not printers)*
Prince & Lemon Sts., Lancaster, Lancaster, Pa. 17604

5. LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS *(Not printers)*
1200 Seventeenth St., N. W., Washington, D. C. 20036

6. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR

PUBLISHER *(Name and address)*
American Psychological Association, 1200 - 17th Street, N.W., Washington, D.C. 20036

EDITOR *(Name and address)*
George Mandler, South Parks Road, Oxford, OX13PS, England

MANAGER EDITOR *(Name and address)*
Harold P. Van Cott, 1200 - 17th Street, N.W., Washington, D.C. 20036

7. OWNER *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)*

NAME	ADDRESS
American Psychological Association	1200 - 17th Street, N.W., Washington, D.C.20036

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES *(If there are none, so state)*

NAME	ADDRESS
None	

9. FOR OPTIONAL COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES (Section 132.121, Postal Service Manual)

39 U. S. C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates."

In accordance with the provisions of this statute, I hereby request permission to mail the publication named in Item 1 at the reduced postage rates presently authorized by 39 U. S. C. 3626.

(Signature and title of editor, publisher, business manager, or owner)

10. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES *(Section 132.122, Postal Manual)*

(Check one)

The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes ☒ Have not changed during preceding 12 months ☐ Have changed during preceding 12 months *(If changed, publisher must submit explanation of change with this statement.)*

11. EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NUMBER OF COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A. TOTAL NO. COPIES PRINTED *(Net Press Run)*	12,102	11,659
B. PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	None	None
2. MAIL SUBSCRIPTIONS	10,186	10,557
C. TOTAL PAID CIRCULATION	10,186	10,557
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS 1. SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	128	128
2. COPIES DISTRIBUTED TO NEWS AGENTS, BUT NOT SOLD	None	None
E. TOTAL DISTRIBUTION *(Sum of C and D)*	10,314	10,685
F. OFFICE USE, LEFT-OVER, UNACCOUNTED, SPOILED AFTER PRINTING	1,788	974
G. TOTAL *(Sum of E & F–should equal net press run shown in A)*	12,102	11,659

I certify that the statements made by me above are correct and complete.

▶ *(Signature of editor, publisher, business manager, or owner)*
Henry S. Salmon Jr Bus Mgr.

PS Form 3526 July 1971

2 FEB 1974

DM
1/24

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 79, No. 2
March 1972

Articles

Recognition and Retrieval Processes in Free Recall **John R. Anderson and Gordon H. Bower** 97

Preperceptual Images, Processing Time, and Perceptual Units in Auditory Perception **Dominic W. Massaro** 124

A Modified Belief Theory of Prejudice Emphasizing the Mutual Causality of Racial Prejudice and Anticipated Belief Differences . . . . **Richard A. Dienstbier** 146

A Stimulus-Sampling Model of the Partial Reinforcement Effect . . . . **Ronald L. Koteskey** 161

Theoretical Notes

Can a Two-State Model Account for Two-Choice Reaction-Time Data? . . . . . . **John Theios and Peter G. Smith** 172

Continuity Theory Revisited: Reply to D. B. Berch **Gerald B. Biederman** 178

Comments on "Roles of Activation and Inhibition in Sex Differences in Cognitive Abilities" by D. M. Broverman, E. L. Klaiber, Y. Kobayshi, and W. Vogel **Mary B. Parlee** 180

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

LEONARD BERKOWITZ
E. J. CAPALDI
R. G. CROWDER
C. R. GALLISTEL
R. GELMAN
R. GLASER
JAMES GREENO
JAMES V. HINRICHS
D. HOWES
MURRAY E. JARVIK
H. KELLY
G. KEPPEL
WALTER KINTSCH
T. K. LANDAUER
GERLAD S. LEVENTHAL
EDWIN MARTIN
E. J. MURRAY
RICHARD M. SHIFFRIN
JAMES T. TEDESCHI
JOHN THEIOS
JOHN THIBAUT
R. E. ULRICH
P. VITZ
R. M. WARREN
RUTH WYLIE

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts submitted to the *Psychological Review*. In each issue the Editor will list in the *Review* those editorial consultants who have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowledged.)

The *Psychological Review* is devoted to articles of theoretical significance to any area of scientific endeavor in psychology. Except for occasional articles, manuscripts exceeding 7,500 words will be discouraged. Ordinarily, manuscripts which consist primarily of original reports of research or of reviews of the literature should be submitted to other journals. Methodological and statistical articles are not appropriate for the *Psychological Review*. See the editorials in the *Review* (1965, **72**, 1–2, 1971, **78**, 1–2) for further details.

Beginning with the 1971 issue, Letters to the Editor will be published on an experimental basis. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obtained from the APA Central Office.

All manuscripts must be submitted in triplicate. In the reference list, give journal titles in full; do not abbreviate. Figures should be submitted as glossy prints; duplicate figures may be photographic or pencil-drawn copies. Authors should retain a copy of the manuscript in case there is loss in the mail.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Manuscripts should be addressed to the Editor, George Mandler, Department of Experimental Psychology, South Parks Road, Oxford OX1 3PS, England.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of materials copyrighted by the American Psychological Association: Written permission must be obtained for copying or reprinting any tables, figures, or text of more than 500 words in length. Persons or organizations who obtain permission to reproduce any such copyrighted materials must include on the first page of the reproduced materials the exact copyright notice which appeared on the Association's original publication. There will be a charge of $10.00 per page for such permission except when the requester is the author of the article to be reprinted, or when the article is to be reproduced in a limited number of copies solely for instructional purposes.

HAROLD P. VAN COTT
Managing Editor

LORRAINE BIROS
Technical
Editor

HENRY S. SALMON
Advertising and Business
Manager

ANNE REDMAN
Subscription Manager

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1200 Seventeenth St., N. W., Washington, D. C. 20036

$10.00 volume

Second-class postage paid at Lancaster, Pa.

$2.00 issue

Send all communications, including address changes, to 1200 Seventeenth St., N. W., Washington, D. C. 20036. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PSYCHOLOGICAL
REVIEW

Diary No. 727(6)
Date 1/2
.. N : 66
ureau Ed. Psy. Research.

VOL. 79, NO. 2 MARCH 1972

RECOGNITION AND RETRIEVAL PROCESSES IN FREE RECALL[1]

JOHN R. ANDERSON[2] AND GORDON H. BOWER

Stanford University

A model of free recall is described which identifies two processes in free recall: a retrieval process by which the subject accesses the words, and a recognition process by which the subject decides whether an implicitly retrieved word is a to-be-recalled word. Submodels for the recognition process and the retrieval process are described. The recognition model assumes that during the study phase, the subject associates "list markers" to the to-be-recalled words. The establishment of such associates is postulated to be an all-or-none stochastic process. In the test phase, the subject recognizes to-be-recalled words by deciding which words have relevant list markers as associates. A signal detectability model is developed for this decision process. The retrieval model is introduced as a computer program that tags associative paths between list words. In several experiments, subjects studied and were tested on a sequence of overlapping sublists sampled from a master set of common nouns. The two-process model predicts that the subject's ability to retrieve the words should increase as more overlapping sublists are studied, but his ability to differentiate the words on the most recent list should deteriorate. Experiments confirmed this predicted dissociation of recognition and retrieval. Further predictions derived from the free recall model were also supported.

This paper has several aims: First, we offer a critique of two popular "strength" theories, one which relates recall and recognition to the strength of one and the same memory trace, and another which relates only recognition memory to a similar strength measure; second, we develop a particular conceptualization about recognition memory which we believe satisfies the criticisms of the traditional strength theory; third, we illustrate how that recognition mechanism could be interfaced with a retrieval mechanism so as to yield a viable theory about multilist free recall. Along with reviewing published data relevant to these points, we also shall present new experimental evidence bearing on the second and third points.

TWO THEORIES OF RECOGNITION AND RECALL

Kintsch (1970) provides a careful review of the strength theory of recall and recognition which supposes that recall and recognition involve basically the same process, except that recognition of an item requires a lower "threshold of strength" than does recall. It is because of this lower threshold that recognition of an item is easier than recall. This strength theory has been an important learning theory for years (cf. Bahrick, 1965; McDougall, 1904; Postman, 1963). Support for this theory comes from the fact that several experimental variables affect recall and recognition in the same way. For instance, temporal variables such as study time, the retention interval,

---

[1] This research was supported by Grant MH-13950-04 from the National Institute of Mental Health to Gordon Bower.

[2] Requests for reprints should be sent to John Anderson, Department of Psychology, Stanford University, Stanford, California 94305.

© 1972 by the American Psychological Association, Inc.

and massing and spacing of presentations have similar effects on recognition and recall (Kintsch, 1966; Olson, 1969). Also, the two tasks display similar serial position curves with small primacy and large recency effects (Shiffrin, 1970b).

However, in contrast to these communalities, there are a number of variables that affect recall and recognition in different ways, suggesting that different processes underlie the two performances. As a first example, words of low frequency in the Thorndike-Lorge count of text are more easily recognized (Schwartz & Rouse, 1961; Shepard, 1967), but are recalled more poorly than words of high frequency (Hall, 1954). As a second example, subjects learning under intentional instructions recall better than subjects learning under incidental instructions, but this ordering of the groups is actually reversed for recognition (Dornbush & Winnich, 1964; Eagle & Leiter, 1964; Postman, Adams, & Phillips, 1955). Finally, associative and categorical relationships among items play an important role in recall, but appear to have little if any effect on recognition (Cofer, 1967; Kintsch, 1966).

An alternate theory about the relation between recognition and recall which has as prestigious a history as the strength theory is the two-process theory of recall (James, 1890; Kintsch, 1970; Müller, 1913). This asserts that while free recall involves processes similar to those occurring in recognition, recall differs qualitatively in that additional search processes are involved. Specifically, it is assumed that during recall, an item is first retrieved from memory by the search process (whatever that is), and then it is tested by the recognition process to determine if it is from the to-be-recalled list. Thus, in order for a word to be recalled it must both be successfully retrieved and recognized. Within this two-process model of recall, Kintsch was able to give a natural interpretation of the experimental variables which differentially affect recall and recognition. However, these interpretations were clearly post hoc, and therefore, as Kintsch (1970) himself admits, "little direct evidence for the two-stage theory is available at this time [p. 341]." One of the purposes of the experiments to be reported in this paper is to provide some evidence directly bearing on this issue.

STRENGTH THEORIES OF RECOGNITION

Current literature on memory appears to present a solid concensus regarding the dominant theory of recognition memory—a theory which we will criticize here and to which we will offer an alternative. The dominant theory assumes that upon presentation of a stimulus on a recognition-memory test, the subject can access directly a representation or trace of that item in memory. Stored with this representation is some single continuous (or many-valued discrete) variable that provides a measure of the subject's degree or strength of familiarity with the test item. We will call this measure the "strength" of the item, using the term formally introduced by Wickelgren and Norman (1966). However, this measure has been given many verbal guises—"familiarity" by Kintsch (1970) and Parks (1966); "amount of information" by Freund, Loftus, and Atkinson (1969); and "number of feature matches" by Bower (1967). The important feature to note about the strength theory of recognition is that it is "ahistorical"; that is, it assumes that a subject makes recognition decisions about an item not on the basis of detailed memory of the past history of occurrences of the item, but rather on the basis of a single measure which reflects to some extent its past frequency, recency, and duration of exposure. It is this ahistorical character of strength theory which is the source of all its weaknesses.

The strength conception of recognition memory is practically as old as experimental psychology (see e.g., the use of "excitatory potential" by Hull, Hovland, Ross, Hall, Perkins, & Fitch, 1940). However, a significant technical advance in strength theory came with Egan's (1958) application of statistical decision theory to recognition memory. That approach assumes that there is one normal distribution of strengths for those items which have been studied recently, and a different normal distribution for those items which have not. The first distribution is called the "old" and the second the "new." The average distance between the two distributions reflects the amount of strength that was added to the studied items by their recent presentation. It is assumed that the subject chooses some criterion point, *C*, along the dimension of strength. (This criterion corresponds to the former "threshold of recognition.") If the strength of a particular item exceeds *C*, the subject will decide it is an old item and thus "recognize" it. If the strength is less than *C*, the subject will decide it is a new item and reject it.

This strength theory, elaborated with the technical machinery of statistical decision theory, can handle some of the salient facts gleaned from simple recognition experiments. For example, it can handle the data taken from experiments in which the subject is asked to rate the confidence of his recognition judgment (e.g., Murdock, 1965). In particular, the theory predicts the bowed-shape memory-operating characteristic relating changes in confidence rating to changes in the proportions of old and new items assigned to the confidence-rating categories. The theory predicts how this memory-operating characteristic changes as a function of retention variables such as recency and frequency (see Kintsch & Carlson, 1967). The theory can also predict the relationship between multiple-choice and single-item recognition experiments (e.g., Green & Moses, 1966; Kintsch, 1968a).

However, two detracting points need to be emphasized regarding these successes. First, these several phenomena are not strong evidence for the theory because the results have been shown to also be quite consistent with several alternate theories of decision making. For example, Kintsch (1968a) showed that the relationship between yes–no and multiple-choice recognition could be well predicted by at least three theories other than statistical decision theory. Similarly, Lockhart and Murdock (1970) have shown that bowed memory-operating-characteristic curves are derivable from a diverse variety of underlying old and new distributions other than the presumed normal distributions. Second, and of greater consequence, these successes are really triumphs (such as they are) for the technical machinery of statistical decision theory, not for the basic strength assumptions. In fact, they provide no discriminating evidence in favor of the strength assumptions. For instance, Bernbach (1967), using a different conception of the underlying memory representation but similar decision-making machinery, produced a recognition model equally as adequate as that generated from strength theory. Similarly, we too will produce an alternate memory model that is equally capable of handling these elementary facts about recognition. The thrust of this argument is that results from simple recognition experiments do not suffice to discriminate among alternative conceptions of the memory representation.

The strength approach can explain our ability to differentiate between a positive and negative set of items on a retention test if these differ in strength; that is, if the items differ in recency, frequency, and duration of study. The approach fails, however, to distinguish theoretically among items which are equated on frequency and recency, but are nonetheless differentiable empirically on other grounds. Yet there are multiple dimensions along which to differentiate sets of items besides their frequency, recency, duration, or some composite strength measure. Example differentia in verbal learning experiments would be *where* in space the item was presented, *who* said it, *how* it was said, and other special characteristics of its physical and psychological context of presentation. In fairness, the original strength theorists never actually said that subjects could *not* perform discriminations on the basis of these dimensions. They were rather concerned with other matters. Strength of familiarity was adopted as a convenient concept for integrating with their technical machinery. However, the evidence is now available that an undifferentiated strength of familiarity concept is not sufficiently rich to account for the subject's ability to differentiate sets of items.

Of relevance to this issue is evidence showing item differentiation in the subject's editing of his free recall. An experiment by Bower and Clark (1969) illustrates differentiation in recall of a sort that cannot be handled in simple strength terms. Their subjects learned lists of unrelated nouns by generating meaningful sentences, woven into stories around the critical words. This heuristic greatly enhanced their later recall of the critical nouns. But, for present purposes, the significant fact is that subjects almost never intruded their elaborative additions while overtly recalling (in writing) the critical nouns. In recalling, they would recite their story to themselves but write down only the critical nouns. That outcome is not an isolated incident. When asked, most subjects in free recall experiments will report thinking about many nonlist words (e.g., a category name) to help mediate recall of list words, yet the mediators are rarely intruded in recall. An experiment by G. H. Bower and P. Winchester (unpublished) showed that following a lengthy word association task, subjects had over 90% accuracy in later identifying whether a given item was one of the experimenter's stimulus words or one of their own associate responses.

The argument above claims that an elemental strength measure provides no grounds for discriminating list items from implicit associative responses during list learning. A related difficulty with a simple strength theory is that it cannot accommodate facts about list differentiation. This issue is engaged, for example, in single-trial free recall studies using multiple lists studied and recalled in succession. In the typical task, the subject is required to recall only the "most recent" list. Now, recall of the most recent list could be implemented in strength models by assuming a retrieval mechanism that outputs only those items whose strength exceeds a criterion amount (earlier items having decayed below that amount). However, as Shiffrin (1970a) has demonstrated, subjects are also capable of recalling only those items from the list which just preceded the most recent list; following presentation of List n, his subjects recalled List $n - 1$, and they could do so with reasonable accuracy.

One might imagine that strength theory could also explain Shiffrin's results by invoking the idea of the decay of memory strength (see Wickelgren & Norman, 1966). Since an item's strength decays continuously between its presentation and its retention test, one might suppose that access to List $n - 1$ is provided by an "internal editor" which recalls only those items whose strengths fall in an appropriate bandwidth (see Hinrichs, 1970), above a lower bound (distinguishing List $n - 1$ from List $n - 2$) and below an upper bound (distinguishing List $n - 1$ from List n). This "bandwidth" explanation receives some support in research by Hintzman and Waters (1969, 1970) showing that identification of which of two lists an item appeared in improves with increasing temporal interval between the two lists but deteriorates with increasing interval between study of the lists and the retention test. So, temporal discriminations (implemented by whatever mechanism) are clearly important in list differentiation.

In counterargument, however, it can be shown that such temporal judgments cannot be based solely on any simple strength variable. This can be concluded from experiments which simultaneously vary frequency and recency. An item in List 1 may be presented 10 times as frequently as an item in List 2; so, according to any "strength" theory, the List 1 item will be the stronger. Yet in List 2 recall, the subject is much more likely to recall the List 2 item than the frequent List 1 item. In studies of list differentiation, Winograd (1968) found in fact that such unbalanced frequencies actually *improved* list discrimination rather than the converse as expected by strength theory. Hintzman and Block (1971) showed that correct assignment of an item to List 1 rather than List 2 increased directly with its frequency of occurrence in List 1; a simple strength theory of list differentiation would have predicted the opposite, namely, higher probability of assigning frequent items to the more recent List 2.

Although list differentiation experiments commonly use each item in just one of several lists presented (see, however, Hintzman & Block, 1971), interesting results are also yielded by experiments in which a given item appears in several lists. We have conducted such an experiment (G. H. Bower & J. R. Anderson, in preparation) in which subjects studied a series of four overlapping lists. There are 16 (2^4) possible combinations of appearances and nonappearances for a particular word across the four lists. Several words were used to instantiate all 16 possible combinations, and subjects were later asked on a retention test to indicate in which lists each word had occurred. For present purposes, the most important result of that experiment is that subjects were much more accurate in list identifications than one could ever predict from a simple strength theory. For instance, subjects can correctly remember that one item occurred in Lists 1 and 4 and that another occurred in Lists 2 and 3, etc.; yet it is simply unimaginable how a single strength measure attached to each item could provide for such patterned discriminations. Such results show that "memory strength" alone is insufficient to account for the salient facts about list discrimination.

To pursue matters just one more countermove, a strength theorist might try to explain such embarrassing results by supplementing the strength information with a further independent measure of the temporal recency of the item. Such a "two-measure" theory might integrate the data in our "2^4" list discrimination experiment along with Winograd's results and those of Hintzman and Waters. However, we would then argue that two measures are still not enough. For instance, the proposed two measures (strength plus a recency tag) cannot explain the ability of subjects to discriminate implicit associates from test items, or their ability to identify the location or the modality (auditory or visual) in which an item was presented (D. L. Hintzman, unpublished). On these grounds it would seem reasonable to suppose that subjects can perform list discrimination on any reasonably salient dimension.

RECOGNITION VIA ASSOCIATION TO CONTEXT

For such reasons, we reject the view that recognition is mediated by a simple strength measure or any simple combination of measures. In place of this we propose that items are recognized by retrieval of certain kinds of context information originally stored along with the item in question. This contextual information includes physical characteristics of an item's presentation, implicit associations to the item, and some cognitive elements representing the list in question. We have no startling insights regarding the nature of these context elements which are presumed to vary across lists. They might include the subject's general mood or attitude, his physical posture, and his physiological state, as well as any conspicuous external cues prevailing during presentation of List n. One prominent set of cues may be provided by implicit verbalizations, especially an implicit count as when the subject says to himself "first list," "second list." Another prominent set of contextual cues for the recognition of an item could be provided by other words in the list. An interesting curiosity is how the subject identifies where List $n - 1$ leaves off and List n begins; but clearly, temporal grouping and distinctive signals, instructions, or activities (like recall) interpolated between lists all combine to promote identification of beginning and end segments of that serial unit we call List n. A further curiosity, not explored here, is the issue of how the subject decides whether temporally distributed clusters of presentations constitute different presentations of the "same" list or are "different" lists.

A significant consequence of this context-retrieval theory is that the task of item recognition is not really different in kind from that of list discrimination. In both tasks, the subject must make his decisions on the basis of contextual information retrievable from the test word. The only difference is that the recognition task involves a yes–no decision; that is, the subject must indicate whether he saw the word in a particular list context. In contrast, list discrimination involves a forced-choice task; that is, the subject must decide in which list context the word occurred.

At this point, however, the proponent of the strength of familiarity hypothesis might want to object and insist on a theoretical separation of the simple recognition paradigm from the list discrimination paradigm. His argument might go as follows: "I concede that list discrimination tasks may require something beyond a strength measure, perhaps something like your 'contextual associations' business. But that does not disprove my strength theory of pure recognition. In that task, surely, subjects use strength of familiarity because it is convenient and because it works. In other words, I claim that simple recognition and list discrimination involve fundamentally different mnemonic and decision processes."

As long as both models handle equally well the simple recognition task, there is no way to disprove this dualistic claim. However, on grounds of parsimony one would say that there is no need for a special theory for a special circumstance when a more general theory subsumes the special circumstance and many others. Besides its lack of parsimony, this attempt to give a special status to the pure recognition paradigm ignores the fact that all recognition experiments are implicitly list or set discrimination experiments. This is obvious upon closer analysis of the "pure recognition" experiments. Every subject has seen, heard, and thought thousands of words before he sees the experimental list, and he frequently sees, hears, and thinks of a great many nonlist words between study and test. We have already noted that subjects think implicitly of many words while the study list is being presented. Undoubtedly, included among the distractors in a "pure recognition" test will be words that the subject encountered before, during, and after the study list in question. Therefore, the subject's actual task is to discriminate a particular subset of words from all others on the basis of differentia such as where, when, and how the word was encountered.

For example, in a recent experiment of ours (G. H. Bower and J. R. Anderson, in preparation), the implicit discrimination involved in an alleged "pure recognition" experiment was explicitly given experimental analysis. After hearing a list of study words which they were told to remember, the subjects heard lengthy instructions describing the multiple-choice recognition test they were about to receive. The multiple choice included one word from the study list along with two distractors. Unknown to the subject, however, one of the distractors was a semantically similar word that had been mentioned two times in the instructions interpolated between the study list and the recognition retention test. In comparison to the study-list item, this "instruction distractor" should have been greatly favored in terms of a strength measure, because it enjoyed both higher frequency and shorter recency than the study-list word. However, as intuition and our model suggest, the subjects had not the slightest difficulty in selecting the study-list word of this multiple-choice set. Their ability to discriminate was also unrelated to whether the subject indicated any awareness of the instructional distractors. The instructional distractor was chosen just slightly more than the totally new distractor, a result consonant with some generalization among the "list" and "instructional" contexts associated to the words. This small demonstration just makes transparent the point we have been urging: What is, to appearances, a simple recognition memory experiment is, in actuality, a list differentiation experiment in disguise.

DETAILS OF THE MNEMONIC REPRESENTATIONS

We could represent the entire constellation of contextual stimuli prevailing during List *n* presentation as a single pattern which is treated in the learning system as a single unit, subject to all-or-none association to items appearing on List *n*. However, for technical and theoretical reasons, it is desirable to postulate a pool of hypothetical elements which singly, separately, and independently serve to identify List *n*. As in the stimulus-sampling theory of Estes (1959), List *n* will then be represented as a population of a number of variable stimulus components. There will be one such stimulus population corresponding to each ordinal list—one for List 1, one for List 2, and so on. It would be realistic to assume that the successive stimulus populations form an ordered array of uniformly overlapping sets so that stimulus generalization, confusion, or failure of list differentiation would be most likely among temporally adjacent lists. However, the basis for that proximity metric will not be presented here, but it has been developed elsewhere (Bower, 1972b).

Our basic conception of human memory is that of a huge network of tens of thousands of nodes interconnected by associations. The nodes correspond to individual concepts the subject has, and the associations encode relations that the subject has learned between the concepts. Upon presentation of a word in study, we assume that the sensory features of that word activate the node in the network that corresponds to the word. Simultaneously, there are active in the network nodes corresponding to the various contextual stimuli that the subject is attending to. We will make two specific assumptions with respect to how the node corresponding to the word becomes connected to the contextual nodes:

1. At the time of occurrence of any word in the list, we suppose that there is simultaneously active a unique element or node in memory which will be called the "list marker" or "list tag." The purpose of this list marker is to record the context prevailing during that presentation of the word. It does this by interconnecting the set of contextual nodes active at that moment. The marker thus acts like a label for a collection or bundle of contextual elements. We assume that there is a probability θ that any particular element in the List *n* population of context elements is active and is associated to the list marker.

2. There is a probability α that the subject will form an association between the memory node corresponding to the word and the memory node corresponding to the list marker. This association serves to record a particular fact about the word, namely, that it occurred in a particular experimental context. Thus, we assume, in common with Mandler (1967), Kintsch (1970), Norman and Rumelhart (1970), and many others, that what is learned in a verbal learning task is not the word itself but rather information about the word (e.g., that it was presented in surprising red letters in the first part of the list while the subject was thinking about what he would buy with his wages from service in the experiment).

Figure 1 provides a schematic diagram of what we have been discussing. Three words in memory, A, B, and C are being associated to a set of contextual elements from List *i* and a set from List *j*. The figure illustrates some overlap between the sets of elements for Lists *i* and *j* because, in general, there is reason to believe that the contextual elements for a particular list will not all be unique. These overlapping elements do not serve to discriminate List

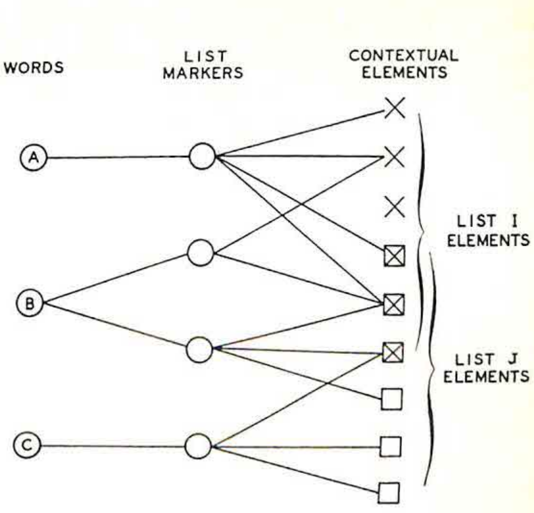


FIG. 1. A hypothetical state of marking of three words A, B, and C, with respect to two lists *i* and *j*.

from List *j*, but they could help discriminate both lists from a third list. In Figure 1, Word A has been successfully associated to a List *i* marker, Word B to both a List *i* and a List *j* marker, and Word C to just a List *j* marker. The list markers serve to keep track of the lists in which a particular word occurred.

One may ask why we have inserted, in Figure 1, list markers between the contextual elements and the words. Why not simply assume that the subject associates the contextual elements directly to the words and eliminate list markers entirely? The reader should keep in mind that the point of introducing associations to context is to provide a means for keeping track of the occasions and the lists in which particular words appeared. This would be difficult to implement on the basis of direct associations between the word and the contextual elements. How would the subject sort out which contextual elements belonged to which list? But even greater difficulties arise for this alternate hypothesis when we consider the problem of how a subject would keep track of what happens within a single list. For instance, if a word appeared twice in a single list, we expect that a subject could often report that fact and describe the contexts prevailing at each of the presentations. According to the model in Figure 1, the subject would have two distinct list markers associated to the word. Each of these list markers would have associated to it a different subset of the contextual elements. Thus, when tested, the subject could retrieve two distinct list markers and therefore judge that he had seen the word in two contexts. Moreover, by means of the contextual elements associated to each of the list markers, he might be able to describe something about the two contexts. For instance, Hintzman and Block (1971) showed that the subjects had considerable accuracy in reporting both serial positions of a word that occurred twice in one list. On the other hand, consider the predicament of the subject if all he had was direct word to context-element associations. If the word was presented twice, more contextual elements would be associated to the word, so the subject might be able to judge that the word had been presented twice. However, he would have no way of sorting out which contextual elements belonged to which presentation. To reiterate, the point of introducing list markers in an association theory is to give the subject a means of keeping different events distinct in his memory.

The notion of list markers developed above is very similar to a model suggested by Hintzman and Block (1971). They also briefly considered an alternate model in which item repetition is reflected by multiple copies or replicas of each word (see also Bernbach, 1969; Bower, 1967). Within such a model, the contextual elements would be directly associated to the word without an intervening marker label, but to a different replica of the word each time it was presented. This approach may appear to eliminate the need for the list markers and the association between the words and the list markers. However, that appearance is deceiving. In such a multiple-copy model, it is necessary to postulate some means of retrieving the many copies from a single phonemic representation of the word (e.g., in order to estimate its frequency of presentation in a list). When such a model is elaborated with the necessary retrieval mechanism, it turns out to be isomorphic to ours: The prototypical phonemic representation of the word corresponds to the word in Figure 1; the different copies correspond to the different list markers; and the probability of retrieving a copy from a phonemic representation corresponds to the probability of forming an association between the word and a list marker.

FORMAL MODEL FOR RECOGNITION JUDGMENTS

The foregoing assumptions provide us with a formal model for recognition judgments that is indistinguishable from Bernbach's (1967). In deciding whether a test word was in List *n*, the subject evaluates the item with respect to how much evidence there is for the word's membership in List n. This is done by assessing the number of List n elements associated to the list marker. During study of any word in List n, there is probability α that a List n marker is associated to the word (α would be assumed to vary with study-time and similar "strengthening" variables). If a List n marker is associated to the word, the number of List n elements associated to the node of that marker will be binomially distributed with parameters k, the number of elements in the List n population, and θ, the probability that a List n element is associated to the list marker (see our earlier assumptions). For large k, it is well known that a binomial distribution approximates to a normal distribution; so such normal approximations will be used throughout this paper. We will let $f_m(x)$ denote this probability distribution of x, the amount of evidence toward List n, for a test item that has been successfully tagged with a list marker. In the following we will refer for convenience to x as the scale of "List n evidence."

We let $f_u(x)$ denote the comparable distribution of evidence for List n for items *not* presented on List n as well as for items presented on List n but not successfully associated to a List n marker, which event occurs with probability $1 - \alpha$. One basis for $f_u(x)$ having nonzero values could be overlap or generalization between successive list contexts in which the test item occurred, and hence confusions. For such reasons, a test item may be judged falsely as occurring on List n when in fact it occurred on earlier lists.

The distributions $f_m(x)$ and $f_u(x)$, and the parameter α, are precisely those required in Bernbach's model. The probability distribution of list evidence for an item presented in the ith list is the probability mixture

$$f_i(x) = \alpha f_m(x) + (1 - \alpha) f_u(x).$$

The distribution $f_i(x)$ is definitely not normal, depending as it does on α and the means and variances of $f_m(x)$ and $f_u(x)$. Figure 2 illustrates some possible distributions for $f_i(x)$ when f_m and f_u have equal variances.

The ingredients for a "statistical decision theory" analysis of recognition memory are now at hand. The distributions $f_u(x)$ and $f_i(x)$ serve, respectively, like the noise and signal-plus-noise distributions in the theory of signal detectability. So the whole set of facts outlined previously regarding yes–no and multiple-choice recognition procedures and memory-operating characteristics are within explanatory reach of the model. The model also extends in the customary way to handle confidence ratings regarding an item's membership in List n. It is supposed that the subject assigns a rating dependent on how much evidence (number of associated List n elements) relevant to List n membership is retrieved by the test item.

The foregoing remarks develop our model for recognition and list differentiation, and it is directly relevant to such data. However, we are here concerned with the recognition process as a monitor or editor for free recall. Following input of a word list, the cue for recall initiates some search and retrieval mechanisms which begin implicitly producing plausible list candidates, items which might be on List n. These implicit candidates are then input to a recognition process to assess evidence for their List n membership. If that evidence exceeds criterion, the candidate is designated a List n item and it is given in overt recall.

RETRIEVAL PROCESS

Having completed our description of the recognition model, we will now outline our model of the retrieval process which generates implicit candidate words for possible overt recall. It is simply hopeless to imagine a random mechanism which searches haphazardly and serially (or in parallel, for that matter) through all possible memory locations, hoping to stumble across a few of the list words. Our memories are very large and not well organized for such random searches. One obviously needs a *directed* search process, and the model to be described accomplishes this by searching only those associative pathways which have been recently tagged as useful for retrieving the word set under considera-

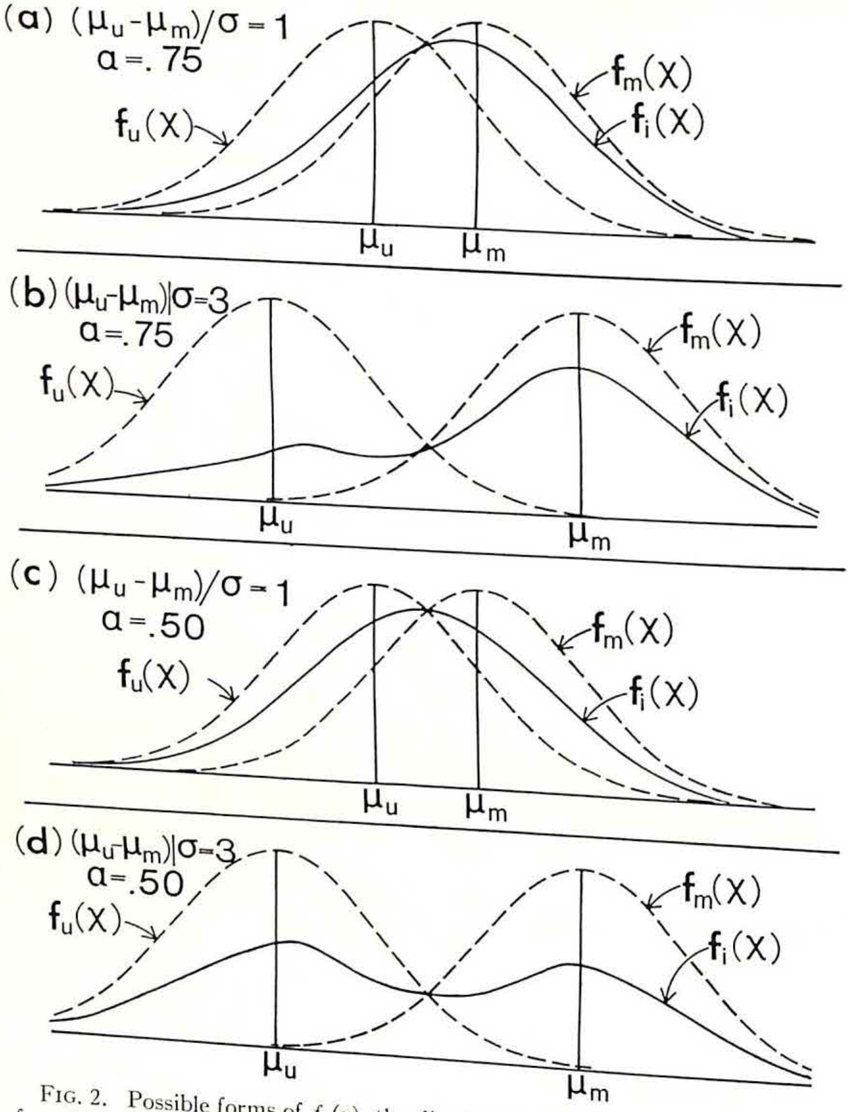


FIG. 2. Possible forms of $f_i(x)$, the distribution of List i evidence for words from List i, as a function of the probability of association to the list marker α, and the distance between the distribution for marked words, $f_m(x)$, and the distribution for unmarked words, $f_u(x)$.

tion. This model is realized in a computer simulation program described by Anderson (1972). Since its particular task environment and testing ground was free recall of word lists, it has been dubbed FRAN, an acronym for Free Recall in an Associative Network.

FRAN starts with a preexperimental network of associations among a large set of concepts (words). The experimenter selects a subset of these words to compose a free recall list; this list is then presented to FRAN one word at a time. FRAN commences to develop a way to recall the words on this presented list (call it List n) and only these. FRAN learns which words are presented on List n by tagging the memory nodes (corresponding to the words being presented) with a List n marker. This corresponds to the recognition process we have just described. In addition to this tagging of the memory node (call it A), FRAN follows one or more associative pathways radiating out from Node A, searching for other words (memory nodes) that are also on the list. This latter information is provided by whether or not the node being examined has a previous association to a List *n* marker. If another list word is found in this associative search, then FRAN attempts to tag the associative pathway from Node A to the found target node. This tagging of a prior associative pathway proceeds just like the tagging of a memory node, namely, by establishing an association to the List *n* marker. These list tags on associative pathways are what will later guide the search process; a tagged pathway essentially delivers a later directive to the search process to follow out this pathway to get from recall of Word A to recall of other list words.

There is one further process of significance occurring during the study phase of a free recall experiment and that is construction of a sublist called ENTRYSET. The list of words serve as special "starters" from which FRAN can begin her chains of recall during output. Technically, a word on ENTRYSET is simply one which receives an association from a special memory node called LIST *N*, which we may consider as a prototypical representation of List *n*. The number of "recall starters" on ENTRYSET is assumed to be very limited (in the current program, only three items). The program has a set of crude heuristics for eventually converging upon those items which are most "central" in the associative network in that they lead to recall of the largest number of list words.

In the usual free recall experiment, during output FRAN first dumps out the three to five recent list words in her short-term memory. She then uses these plus the items on ENTRYSET to start her search of long-term memory for further words to recall. This search involves following the tagged associative pathways radiating out from each of the starter nodes in turn, with each search proceeding in a "depth first" manner. During this directed associative search, FRAN considers many words, testing each for recognition (i.e., whether it has an association to a List *n* marker). If a retrieved word is recognized, it is overtly recalled.

The reader will note that context information is being used in three different ways in FRAN. First, via the list marker (collections of context elements), retrieval of context from a test node underlies correct identification of List *n* words, thus forming the basis for list discrimination. Second, via the LIST *N* node (i.e., the prototype of the List *n* context), the context also has the status of a "stimulus" initiating recall chains from the ENTRYSET words. These two uses of context are not original with us. Similar notions were suggested by Norman and Rumelhart (1970); a standard assumption of interference theory is that context cues serve as stimuli governing emission of a response (Keppel, 1968; McGovern, 1964). We have expanded these notions and provided an explicit representation and model for them. The third use of context, to tag associative paths in the memory network, is unique to our learning and retrieval model. Such second-order associations, that is, an association between an associative link and a memory node (a List *n* marker), have not been used in the more traditional associative theories. However, the second-order association is really only a labeling of a lower order association, and most current theories of memory (e.g., Quillian, 1969; Rumelhart, Lindsay, & Norman, 1972) do involve labeling of associations with such semantic relations as category membership, superordination, opposition, and the like. In FRAN, labeling preexisting associative pathways with a List *n* tag provides the mechanism for guiding FRAN's memory search during recall. Since the two kinds of tagging, of memory nodes versus associative pathways, are independent, it is the case that word recognition in FRAN is independent of the retrieval processes. This independence of the two processes will be important in the experiments that follow.

The above description of FRAN is necessarily brief and incomplete, focused on only those aspects pertinent to the two-process recall hypothesis. Papers by Anderson (1972) and Bower (1972a) should be consulted for further details of the theory, the simulation program, and a sampling of results which the theory explains and those with which it has difficulties. In our opinion, it is the most viable theory of free recall currently available in terms of its range of applicability.

EXPERIMENT I

Experiment I is one of those research curiosities begun for irrelevant reasons but which turned out to provide the stimulus for this program of research and theorizing. We presented subjects 15 successive lists of 16 words, each list selected in a pseudorandom manner from a master set of 32 words. As a consequence, lists after the first had partial overlap in membership with earlier lists. After the presentation of each list, the subjects were required to recall only those words that were in the most recent list. According to our retrieval model, the ability to retrieve the full set of 32 words should improve with repeated study trials because there is more time to discover and tag associative pathways between the words.

However, the model makes a markedly different prediction about the recognition subprocess of free recall. The implicit candidates for recall generated by the retrieval component will include many or most of the master set of 32 plus some associatively related items. The task of the recognition component is to select which of these words occurred in the most recent list. This it does by checking to see which of the candidates are tagged with list markers referring to the most recent list. Since a different set of 16 words must be recalled on each trial, the subject can only succeed if he associates a different list marker with the to-be-recalled words on each trial (list). Tagging with list markers is conceived essentially as a paired-associate task in which the word functions as the stimulus term and the marker functions as the response term. Thus, the tagging of a single word appearing in successive lists exemplifies the classical A–B, A–C paradigm for negative transfer. One might assume that the amount of negative transfer accumulates as the subject learns A–B, then A–C, then A–D, etc. Similarly, we assume negative transfer grows as the subject learns associations to list markers as in A–List 1, A–List 3, A–List 5, and so forth. We have not found a cumulative negative transfer study in the literature that exactly parallels our situation, but Underwood (1945) has demonstrated that proactive inhibition increases as the same stimulus term is paired with progressively more responses. It is reasonable to assume that negative transfer would also increase as the same stimulus is paired with more responses. For instance, such a result is predicted by the assumptions about response competition proposed by Kjeldergaard (1968).

In this experiment, each word will be paired with a new list marker on every trial that it appears for study. Therefore, the probability of successfully tagging a word should decrease across lists, with a consequent decrease in the subject's ability to recognize to-be-recalled words. Thus, a dissociation of the two subprocesses is expected, with retrieval improving and recognition of implicitly retrieved words deteriorating across successive lists. Such a dissociation would emphasize the utility of conceiving of free recall in terms of the two-process model. It would be impossible for the threshold model, which relates both recognition and retrieval to the same factor of "strength," to account for such a dissociation.

Method

From a master set of 32 common concrete nouns, 15 lists of 16 words were generated. Words were assigned randomly to lists with the constraint that the overlap of each list with each other list would be eight words. Such independent, multiple overlapping of lists was done so that the subjects could not use membership of a word in one list as any indication of its membership in any other list. Therefore, decisions about membership of a word in the current list would have to be based solely on the word's rating on the scale of evidence for that list. By the sixth list, every word in the master set had occurred in at least one list. By the fifteenth list, a particular word had occurred in anywhere from 4 to 11 of the lists. The order of words within a particular list was randomly determined. The same sequence of 15 lists was presented to all of the subjects. The words were slide projected one at a time for two seconds. Immediately after presentation of the list, the subjects wrote their recall of the most recent list of words they had just studied. They had 135 seconds for this recall. They were instructed to recall a word only if they thought it "probable" that the word came from the last list studied. The experimenter collected the recall sheets and the procedure was repeated for the next list. Including the instructions, the experimental session lasted about 50 minutes. Seventeen subjects (6 males, 11 females; 16–22 years old) were recruited through an advertisement in a local newspaper and were paid $1.75 for their services. They were tested in two groups of size 7 and 10.

Results

Figure 3 shows the average number of words recalled from the most recent list (abbreviated R words) and separately, the number of words which were intruded from earlier lists and were not on the most recent list (N words). The number of R words recalled minus the number of N words, also shown in Figure 3, provides a conservative correction for guessing in recall. In the curve for R-word recall and in the corrected-recall curve, there appears to be an initial improvement in recall followed by deterioration. Using orthogonal polynomials, an equation involving the first six powers was fitted to the corrected recall over the 15 lists. The improvements in the fit due to the linear and quadratic components of the curve were significant ($F = 8.35$, $df = 1/8$, $p < .05$, for the linear component; $F = 12.36$, $df = 1/8$, $p < .01$, for the quadratic component). However, the addition of the four higher order polynomials did not significantly improve the fit of the theoretical curve ($F = 2.03$, $df = 4/8$, $p > .10$). The smooth curve in Figure 3 describes the best fitting quadratic equation for the corrected recall. There is considerable variability of the observed points about the quadratic curve. This variability may be attributed to differences among the individual word lists. The theoretical interpretation of this overall rise then fall in recall is that the improvement in item retrievability predominated initially, but that after retrievability had reached an upper limit, the degradation of list differentiation continued and recall deteriorated as a consequence. This is because recall is monitored or edited by list differentiation decisions. This interpretation is further tested in the following experiments.

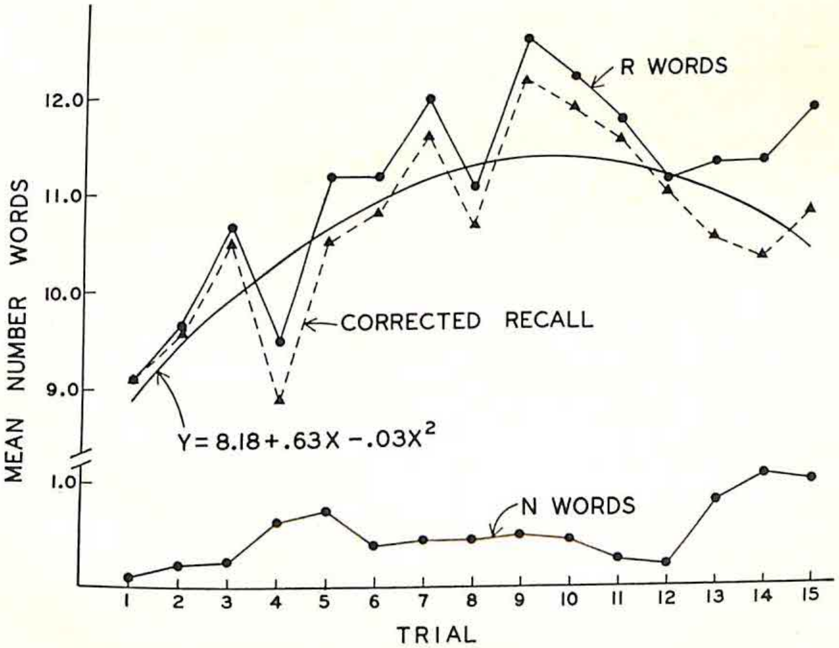


FIG. 3. Mean number of words recalled in Experiment I as a function of trials.

Experiment II

A simple challenge to this interpretation of the results is to question whether the results have anything to do with the particular experimental manipulation employed. Quite possibly the initial improvement may simply reflect "learning to learn" and the later decrement may only reflect progressive fatigue or loss of motivation of the subjects. The obvious control, then, is to repeat the last experiment but with a completely different list (i.e., no words repeated) on each trial. There should be negligible across-trial change in retrievability because the subject is learning new items on every list. Also because new words are being studied on each trial, both the stimuli and the responses for the hypothetical paired-associate task underlying the list marking are different on each trial. Therefore, there should be little, if any, negative transfer in identifying the most recent list. Hence, if the results of Experiment I were really due to the pseudorandom repetition of words across lists, then neither the initial increase nor the later decrease in recall should be observed in this second experiment.

Method

Twenty lists of 16 words were created by sampling randomly without replacement from a master set of 320 common concrete nouns. The same 20 lists were used in the same order for all subjects. The procedure for testing was identical to that in Experiment I. This experiment, including instruction, lasted about 70 minutes. Twenty-one subjects (10 males and 11 females; 18–22 years old) served in this experiment as partial fulfillment of a requirement for the introductory psychology course at Stanford. They were tested in two groups of size 10 and 11.

Results

Figure 4 gives the mean number of words recalled in Experiment II as a function of the trial number. Strictly speaking, only the first 15 trials are relevant to a comparison with Experiment I. Over these trials there is neither a significant linear nor a significant quadratic trend. However, if all 20 trials are considered, the linear trend becomes significant ($F = 8.04$, $df = 1/17$, $p < .05$) although the quadratic trend is still not significant ($F = 1.87$, $df = 1/17$). The smooth curve in Figure 4 describes the best fitting quadratic equation to the data over the 20 trials. The nonsignificant quadratic trend in Figure 4 is exactly opposite to that of Figure 3—that is, recall is worse in the middle trials of the experiment. However, the only significant effect is the linear trend which indicates something of a warm-up effect with recall improving slightly toward the end of the experiment. While the improvement across trials is not very substantial, it contrasts in a minor way with the results of Murdock (1960) who found no improvement across unrelated lists in single-trial free recall.

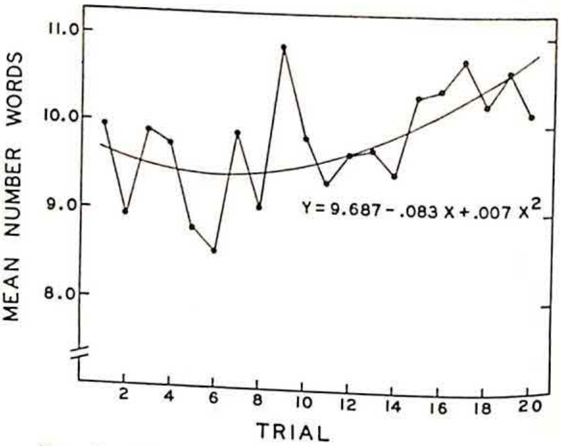


Fig. 4. Mean number of words recalled in Experiment II as a function of trials.

Experiment III

Although the preceding experiments confirmed expectations, the hypothesized mechanisms would be more credible if we could observe separately the decay in recognition and the increase in retrievability rather than viewing only their combined effect on recall. One would then be able to determine whether improvement in retrievability first dominated but was later overcome by deterioration in list recognition. The third experiment provides separate measures of retrieval and list recognition.

Method

The subjects in this experiment studied the identical sequence of 15 lists as used in Experiment I. However, they were required to try to recall the entire master set of 32 words on each trial, not just the 16 presented in the most recent list. The subjects were further required to rate on a 6-point confidence scale whether or not they thought each word they recalled came from the most recent list they had studied. The rating scale ranged from "1 = confident the recalled word was on the most recent list" up to "6 = confident the recalled word was not on the most recent list." Hence, to use the terminology of Experiment I, a low confidence rating indicated that the subject thought the word was an R word, whereas a high confidence rating indicated that he thought it was an N word. By examining recall of R words without regard to their rating, changes in retrievability may be monitored; by examining the confidence ratings, changes in the subject's ability to recognize words can be monitored.

The subjects were given 165 seconds to write their recall and to record confidence ratings beside each of the words recalled. Detailed instructions and illustrations on use of the confidence scale were given since a pilot study had indicated that many subjects would misunderstand the instructions and use low numbers just for those words which appeared for the first time in the most recent list. Including the instructions, the experimental session lasted about 75 minutes. Twenty-four subjects (10 males, 14 females; 16–23 years old) were recruited through an advertisement in a local newspaper. In a two-hour experimental session they participated in this and another unrelated experiment which followed, and they were paid $3.50 for the total session. They were run in three groups of sizes 6, 8, and 10. Six subjects were excluded from the analysis because of failure to use the recognition scale properly on some of the trials.

Results

An initial question is whether the earlier results of Figure 3 have been duplicated under these altered conditions. In this experiment, a rating of 1 indicated that the subject was certain that he had seen the word in the list just studied, and a rating of 2 indicated he thought it probable that the word came from that list. Since the subjects were instructed in Experiment I to recall only the words they thought "probably had occurred" in the last list, the words recalled and rated 1 or 2 in Experiment III should be comparable to those recalled in Experiment I. Figure 5 shows the frequency with which R words (in the most recent list) were recalled and rated 1 or 2, and the frequency with which N words (not in most recent list) were recalled and rated 1 or 2. Figure 5 also shows recall corrected for guessing by subtracting the mean for the N words from the mean for the R words. The data represented in

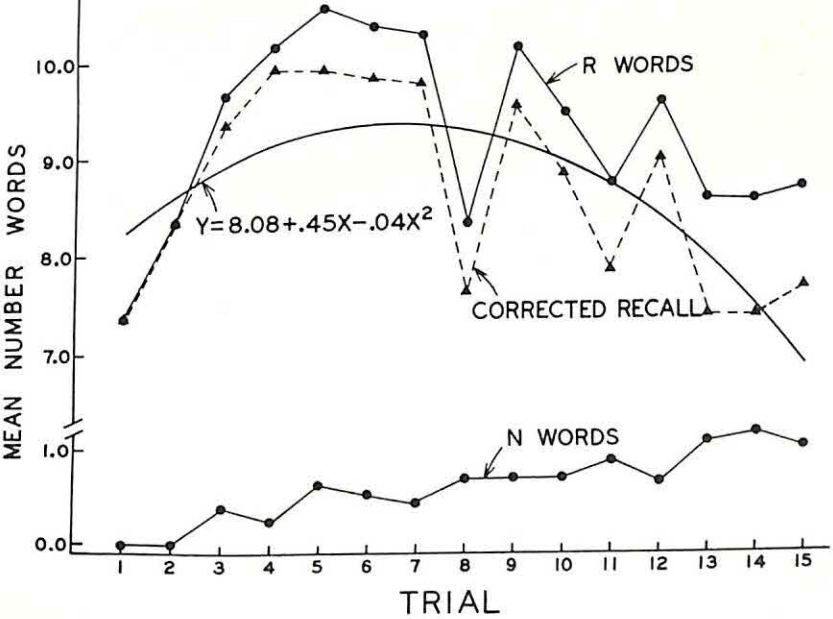


FIG. 5. Mean number of words recalled and rated "1" or "2" in Experiment III as a function of trials.

Figure 3 and the data represented in Figure 5 are quite similar. As in Experiment I, an equation involving the first six powers was fitted by means of orthogonal polynomials to the corrected recall over trials. The improvement in the fit due to the linear component was marginally significant ($F = 5.06$, $df = 1/8$, $p < .10$); the improvement in fit due to the quadratic component as quite significant ($F = 9.54$, $df = 1/8$, $p < .025$); the four higher order polynomials did not significantly improve the fit ($F = .90$, $df = 4/8$). The smooth curve in Figure 5 describes the best fitting quadratic equation to the corrected recall. That quadratic equation confirms an initial rise and subsequent fall in the corrected recall. Thus, we may conclude that the procedural changes in this experiment have not altered the basic processes that were occurring in the first experiment.

We can now determine whether, as hypothesized, retrievability improves while list recognition deteriorates across trials. Figure 6 presents the mean number of words recalled, while Figure 7 shows the mean confidence rating of the words recalled. In these figures, the data are classified into four groups according to whether or not the word was in the most recent list and according to the number of times the word had appeared before that list. The eight R words that had been presented fewer times than the mean number for those in the most recent list were classified as LR (less frequent, recent), and the other eight presented more than the mean, MR (more frequent, recent). When frequencies were tied, the words assigned to the LR group were those for which the lag between the current presentation and the next most recent was the longest. On a similar basis, the N words were subclassified as LN and MN.

In Figure 6 it is clear that recall of R words increases with negative acceleration to an asymptote, confirming the hypothesis about the improvement in retrievability. On the other hand, Figure 7 shows that the mean difference in recognition ratings between R words and N words steadily decreases as a function of trials, confirming the hypothesis about the degradation in list recognition. Thus, the interpretation of the results of Experiment I as arising through the dissociation of retrieval and recognition has been confirmed.

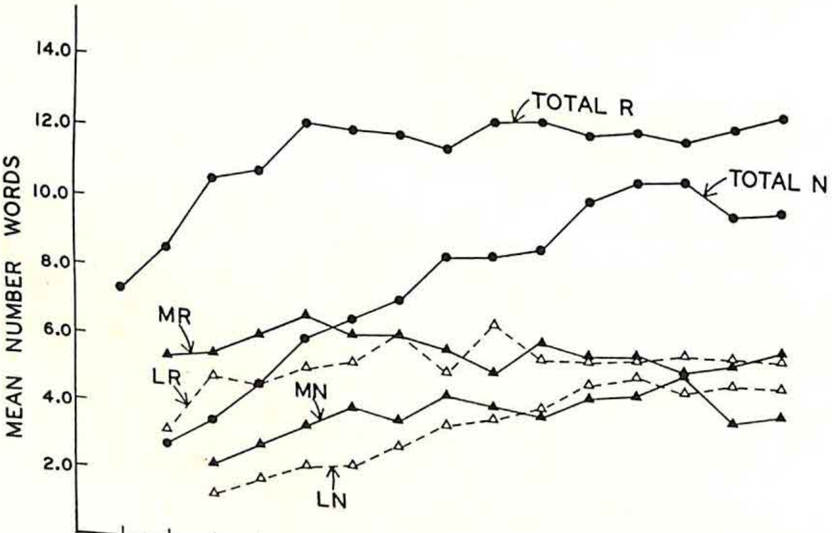


FIG. 6. Mean number of words recalled in Experiment III as a function of trials

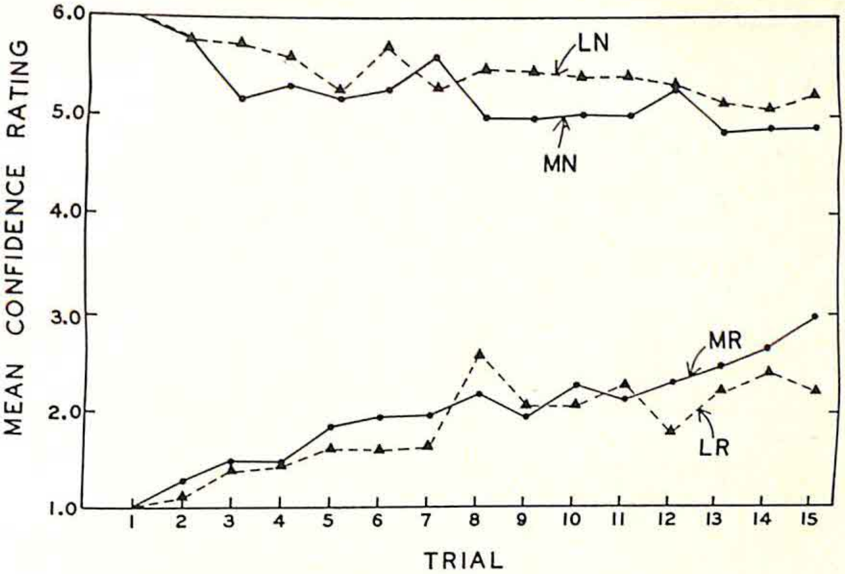


FIG. 7. Mean rating of the words recalled in Experiment III as a function of trials.

Several auxiliary hypotheses arising from our theoretical position may also be examined. First, our model relates improvements in retrieval over trials to the fact that the subject has more opportunities to locate and learn associative paths to access the words. Therefore, those words best recalled should be those that have appeared in the most prior lists—a simple frequency effect. Inspection of Figure 6 confirms that the more frequently occurring words (represented in the curves labeled MR and MN) were initially better recalled than the less frequently occurring words (curves LR and LN). As retrievability approached its asymptotic level, the differences between more and less frequently occurring words diminished, as is to be expected. These conclusions from a visual examination of Figure 6 are fully substantiated by statistical analyses.

Second, on the hypothesis that negative transfer increases as the same word is paired with more list markers, one would predict that the mean confidence rating should be lower (i.e., more correct) for those R words that occurred less frequently. The fewer prior list markers associated to a word, the more "novel" it is, the less the negative transfer in establishing the new "word → List *i*" association. Figure 7 shows 14 trials over which a comparison of confidence ratings can be made between MR and LR words. Of these 14 comparisons, 11 trials exhibit the predicted inequality. To confirm the reliability of this difference, the mean ratings across these 14 trials for the MR and LR words were computed for each of the 18 subjects. The difference in average ratings between the two types of words is highly significant ($t = 4.41$, $df = 17$, $p < .001$). This finding is all the more impressive when one realizes that just the opposite ordering of the MR and LR curves would be predicted on the basis of a strength model which relates recognition to frequency, recency, and duration of exposure.

Considering the data at the top of Figure 7, on 12 of the 13 relevant trials, the less frequent N words were better differentiated (have higher scores) than the more frequent N words. In other words, more frequently presented N words were judged as more likely to have been on the most recent list than were less frequently presented N words. A t test confirms the significance of this difference ($t = 3.31$, $df = 17$, $p < .005$). It is not clear how to interpret this result. Perhaps it could be explained by the mechanism of generalization of contextual elements between lists.

The more prior lists in which an item has appeared, the greater the probability would be that one of the list markers associated would be mistaken as a list marker for the current list. The problem with the generalization mechanism is that it would predict the opposite ordering of the MR and LR curves. To explain why the LR curve is below the MR we would have to make the added and unmotivated assumption that the negative transfer in list tagging outweighed the list generalization factor. In any event, the conclusion to be drawn from Figures 6 and 7 is that, in this paradigm, retrieval is directly and recognition is inversely related to frequency of exposure.

Model Testing

In the introduction, we stated that the distribution of values for nonpresented N words on the scale of List i evidence is the normal $f_u(x)$, but the distribution for the R words is the nonnormal $f_i(x) = \alpha f_m(x) + (1 - \alpha) f_u(x)$. A relatively simple test of this mathematical model is possible. The probability that a particular N word is rated with a confidence exceeding j provides an estimate of the standard normal deviate corresponding to the jth criterion point. In this way, the data from the N words provide estimates of the intervals between the five criterion points. A grid search can then find the probability of tagging, α, and of the mean of the distribution of tagged words, μ_m, that will yield the best fitting distribution $f_i(x)$ for the R words. The best fitting parameters are those that yield predicted frequencies of the confidence ratings that deviate minimally from the observed frequencies as measured by a chi-square test.

One could test the statistical significance of the deviations between observed and predicted frequencies as a goodness-of-fit statistic. However, a few large chi-squares may be expected even if the model were essentially correct because the estimates of the confidence intervals are only approximate and are subject to random error. A fairer test of the model would find the minimum chi-square estimates of seven free parameters, namely, the five criterion points (for the confidence ratings), α, and μ_m. Unfortunately, the computing cost to find seven parameters is many orders of magnitude greater than the cost to estimate two parameters. So, instead of the best absolute test of the model, we will present a comparative test of the model. We will compare our model with a plausible alternative theory, the traditional recognition model that presumes that one normal distribution underlies words from the list and another underlies words not from the list (e.g., Parks, 1966; Wickelgren & Norman, 1966). Since that model assumes that the two distributions have the same variance, the grid search needs only to estimate the mean of the likelihood distribution for R words, d'. The value of d' would characterize a memory-operating characteristic in the traditional signal-detectability analysis.

The test proposed is impossible for Trial 1 because there are no N words recalled; also not enough N words were recalled on Trial 2 to provide reliable estimates of the intervals between the criterion points. Therefore, comparisons of our model with the traditional one will be confined to Trials 3 through 15. Table 1 presents for these trials the estimated values of the criterion points $(c_1, c_2, c_3, c_4, c_5)$, α, μ_m, and d' as well as the chi-square measures for deviations of our model and of the traditional signal-detectability model. For every trial, the predictions of the traditional signal-detectability model lead to large chi-squares, while the assumptions of our model results in smaller chi-squares. Only 4 of the 13 chi-square totals for our model are significant at the .05 level. Notice that as predicted by the principle of negative transfer, there is a decrease across trials in α, the probability of successfully tagging (associating) a word with a list marker. Using the weights suggested by Abelson and Tukey (1963) to test for monotonic trend, the decrease in α across trials is found to be highly significant ($F = 24.32$, $df = 1/11$, $p < .001$). Although less obvious, there is also a marginally significant decrease in the value of

TABLE 1

PARAMETER ESTIMATIONS AND CHI-SQUARE DEVIATIONS FOR THREE MATHEMATICAL MODELS

Trial	Criterian points					Traditional model		Proposed model			Tradition model with variable σi		
	c_1	c_2	c_3	c_4	c_5	d	χ^2 $(df = 4)$	α	μ_m	χ^2 $(df = 3)$	d'	σi	χ^2 $(df = 3)$
3	1.88	1.34	1.18	1.14	.82	2.60	25.39	.91	3.42	.50	4.44	2.38	.80
4	2.33	1.70	1.37	1.08	.61	2.85	54.13	.93	3.63	2.54	4.75	2.40	4.59
5	1.60	1.25	1.13	.90	.51	2.20	32.80	.85	3.08	5.70	3.58	2.34	3.72
6	1.60	1.48	1.37	1.23	.76	2.16	20.99	.95	2.38	16.98	2.68	1.48	15.26
7	1.88	1.55	1.51	1.23	.69	2.34	29.03	.89	2.98	4.79	3.14	1.83	4.26
8	1.88	1.44	1.23	1.08	.47	1.86	36.75	.79	2.54	4.54	2.34	1.82	1.15
9	1.70	1.44	1.10	.93	.37	2.05	46.68	.83	2.81	15.49	3.09	2.14	10.43
10	1.88	1.48	1.10	.86	.40	1.95	62.44	.77	2.91	6.16	2.77	2.17	2.30
11	1.64	1.44	1.18	.82	.47	1.67	36.44	.78	2.37	8.50	2.05	1.64	9.76
12	2.17	1.60	1.20	.92	.45	2.09	103.64	.79	2.99	22.50	2.93	2.31	12.57
13	1.60	1.34	1.04	.64	.11	1.62	35.66	.77	2.34	1.76	2.02	1.72	.66
14	1.44	1.23	.90	.63	.35	1.45	29.34	.70	2.27	5.88	1.79	1.69	4.15
15	1.64	1.34	.97	.66	.40	1.51	52.76	.65	2.69	7.55	1.97	2.15	5.60

μ_m ($F = 6.14$, $df = 1/11$, $p < .05$). This decrease in the distance between the distributions $f_m(x)$ and $f_u(x)$ may reflect increasing confusion between the list marker from the most recent trial and list markers from earlier trials because of the generalization of contextual elements.

Perhaps a more appropriate alternative model would be the traditional signal-detectability theory with the variance of the distribution for R words as a free parameter to be estimated from the data. This model has the advantage of equaling ours in the number of estimated parameters. The fit of this expanded model is shown in the last columns of Table 1, showing the estimates of d' and σ_i, the mean and standard deviation of the $f_i(x)$ distributions. In terms of goodness of fit, there is little basis to choose between this model and ours. This outcome was not entirely unexpected. Inspection of Figure 2 shows that the likelihood distribution $f_i(x)$ is sometimes sufficiently normal to make difficult the discrimination between it and the true normal distribution predicted by the traditional model for the R words. Furthermore, we are dealing with an average of likelihood distributions. There is a different distribution $f_i(x)$ for every word because each word has appeared in a different set of lists and therefore has a different amount of negative transfer associated with it as a stimulus. The average of these many individual distributions will be considerably more normal than any of its constituents.

It is pertinent to examine the changes in the parameters σ_i and d' across trials for this version of the traditional signal-detectability recognition model. The estimates of d', the distance between the means of the distributions, show a fairly consistent decrease across trials. One difficulty with this traditional signal-detectability model is that it provides no theoretical base from which one might predict this decrease in d'. Also, the parameter σ_i is fluctuating erratically from trial to trial. These fluctuations are probably a consequence of the minimum chi-square estimation procedure. Occasionally, events with low theoretical probabilities will have frequencies that are considerably deviant. The estimation procedure gives considerable weight to such deviant low-probability events. In an attempt to minimize deviations of predicted from observed frequency for such an event, the predicted distribution can be markedly altered from one trial to the next.

Discussion of Experiment III

The results of Experiment III showing an improvement in recall alongside a decline in recognition, bring into sharp focus the failing of the "threshold" theory which relates recognition and free recall to a single theoretical construct, such as strength or familiarity. We have dissociated two behavioral measures which that theory claims are coupled by the nature of the system. Such data, along with other arguments advanced in the introduction, render the threshold theory untenable as an account of multilist free recall.

In a similar vein, our results increase the implausibility of a simple strength of familiarity theory of list recognition. The basic problem with such theories is that the subject's judgment of the list membership of an item is not wholly predictable from that item's "composite strength." Both in Experiment I and Experiment III, all 32 words of the master set rapidly became quite "familiar" to the subject. And yet, as Figures 3 and 5 demonstrate, the subjects still showed considerable ability to recognize (discriminate) R words from N words. This discrimination varied only moderately with the frequency of an item's presentation prior to the current list being discriminated.

One might suppose that this R versus N discrimination could be effected by a rapidly decaying "short-term" strength attached to items, such that only items in the most recent list would have strengths exceeding a specific criterion. But this account then finds inexplicable the results of Experiment III in which subjects not only recalled but correctly labeled the N words that were not presented on the most recent list. How are such N items being discriminated from the vast pool of known common nouns in the subject's long-term memory which are not being used at all in the experiment? Such questions, and the several others posed in the introduction, overstrain and discredit the simple strength or familiarity assumptions about list recognition. In place of simple strength, we propose that subjects use something like "list markers" for indexing the lists in which an item has appeared.

One piece of evidence that appears to upset the picture established by the first three experiments comes from a free recall study reported by Ehrlich (1970). He gave his subjects 10 trials of free recall learning on a 20-word list. By the tenth list, recall had reached a near maximum of 19 out of 20. Then the subjects were switched to a series of 10 single-trial free recall trials on semirandom subsets of the original set of 20, much in the manner of Experiment I. Since retrievability for the set of 20 had apparently asymptoted, one might expect to see a continuous decay in the number of words recalled as a consequence of the build-up in negative transfer for list identification. However, the level of recall remained constant over the 10 sublists in Ehrlich's experiment. It is possible that over the initial 10 trials on the whole list in his experiment, negative transfer may have reached its asymptotic level, and hence there would be no more deterioration in the list-tagging process over the last 10 trials on the part lists. By the fifteenth trial in our Experiment I, subjects had seen each word in a mean of 7 lists; by the first part-list trial of his experiment, subjects had seen each word in 10 lists.

The point is that we cannot expect the decay in recognition to continue forever. If our experiment had been extended for more trials, recognition performance would surely decay eventually to some asymptotic level. The two facts that both retrievability increases to an asymptotic level and that recognition also decreases to an asymptote implies that the phenomenon of a rise and fall in recall found in Experiments I and III is in reality a rather delicate matter. It depends on retrievability increasing faster than recognition deteriorates, but asymptoting sooner. In a pilot attempt to replicate Experiment III with altered procedures (more time for recall, different rating scale, different subject population), we replicated the increase in retrieval displayed by Figure 6 and the deterioration of recognition displayed by Figure 7. However, the quadratic trend in Figure 5 (the rise and fall in recall) was relatively slight and in fact nonsignificant statistically.

EXPERIMENT IV

Our model assumes that the processes underlying recognition in free recall are identical to those processes which underlie performance in a pure recognition task. However, evidence has not yet been presented to support this assumption. The next experiment does this by examining pure recognition performance with the same 15 lists that were used in Experiments I and III. If our negative-transfer assumption is correct, we should find the same deterioration in recognition with frequency of presentation as was obtained in Experiment III.

Method

The identical sequence of 15 lists were presented as in Experiments I and III, but with two different testing methods. For the first test method, immediately following the presentation of the list on any trial, the subject had to add mentally 15 single-digit numbers presented at a two-second rate. This manipulation was intended to prevent the subject from recognizing any test word from his short-term memory (see Anderson, 1972). The list marker theory of recognition only applies to items not in short-term memory at the time of testing. In FRAN, any item still in short-term memory at the time of testing is, of course, assured of perfect recognition. After this summation task, the subjects were given a sheet listing the master set of 32 words and they were to rate on a 1 to 8 scale the subjective likelihood or confidence that each word had been on the most recent list. In this experiment, a high rating indicated that the subject thought the word very probably was on the most recent list, and a low number indicated that it very probably was not on the most recent list.

For the second test method, the subject was shown the test items one at a time slide-projected at a five-second rate. The subject rated the items on the same 8-point scale as in the first test method. These two different methods of test were used to insure the generality of the results. The second method corresponds more to how we postulate that recognition happens in free recall—that is, the subject must judge one word at a time. However, there is no compelling reason to expect differences between the two methods. Thirty-one subjects (15 males and 16 females; 18–32 years old) participated in this experiment as partial fulfillment of a requirement for the introductory psychology course at Stanford. They were run in groups ranging in size from 5 to 10. Twenty subjects were tested by the first test method, and 11 by the second test method. Two subjects tested under the first test method were eliminated from the analysis because of incorrect use of the recognition scale.

Results

No systematic differences appeared in the confidence ratings as a function of the method of test, so only the pooled results will be presented. Figure 8 presents the mean confidence ratings for the different types of words across trials. Note that the confidence scale in this experiment was reversed from Experiment III in that in this experiment high numbers mean the subject thought the word was on the most recent list. Also an 8- rather than a 6-point confidence scale was used. It does appear that the mean difference in rating between

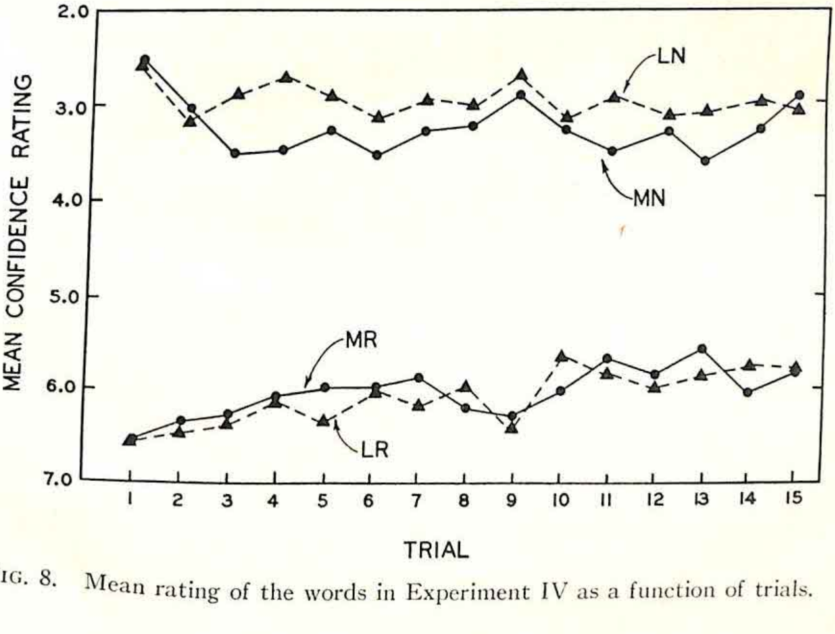


FIG. 8. Mean rating of the words in Experiment IV as a function of trials.

the R words and the N words decreases across trials. On Trials 1–5, the mean difference is 3.31; on Trials 6–10, it is 2.92; on Trials 11–15, it is 2.65. A test for a monotonic decreasing trend in the difference between the R and N words is highly significant ($t = 6.10$, $df = 13$, $p < .001$). However, the deterioration in recognition displayed in Figure 8 is not as dramatic as that in Figure 7 from Experiment III. As in Experiment III, those words presented less frequently (represented in the curves LR and LM) are the more accurately identified. Thus we have replicated the findings of Experiment III with respect to recognition.

From the data of this experiment, we can discover what free recall would be like in an experiment in which retrieval was perfect from the start but recognition was still subject to deterioration through negative transfer in the list-tagging process. Retrieval is perfect because the experimenter provides the subject with all the words; the subject needs only to decide which came from the most recent list. Figure 9 indicates the sort of free recall data that would obtain under the circumstance of perfect retrieval. Here we have the number of R words rated 7 or 8 and the number of N words rated similarly. These may be taken to represent the number recalled and the number intruded in the hypothetical circumstance. We have also plotted in Figure 9 the corrected recall obtained by deducting intrusions from recalls. Figure 9 is to be compared with Figure 3 from Experiment I and Figure 5 from Experiment III. As we would predict, since "retrieval" is asymptotic from the start, there is a rather dramatic decline in "recall" as negative transfer builds up and impairs the recognition component.

Thus, it would appear that we may

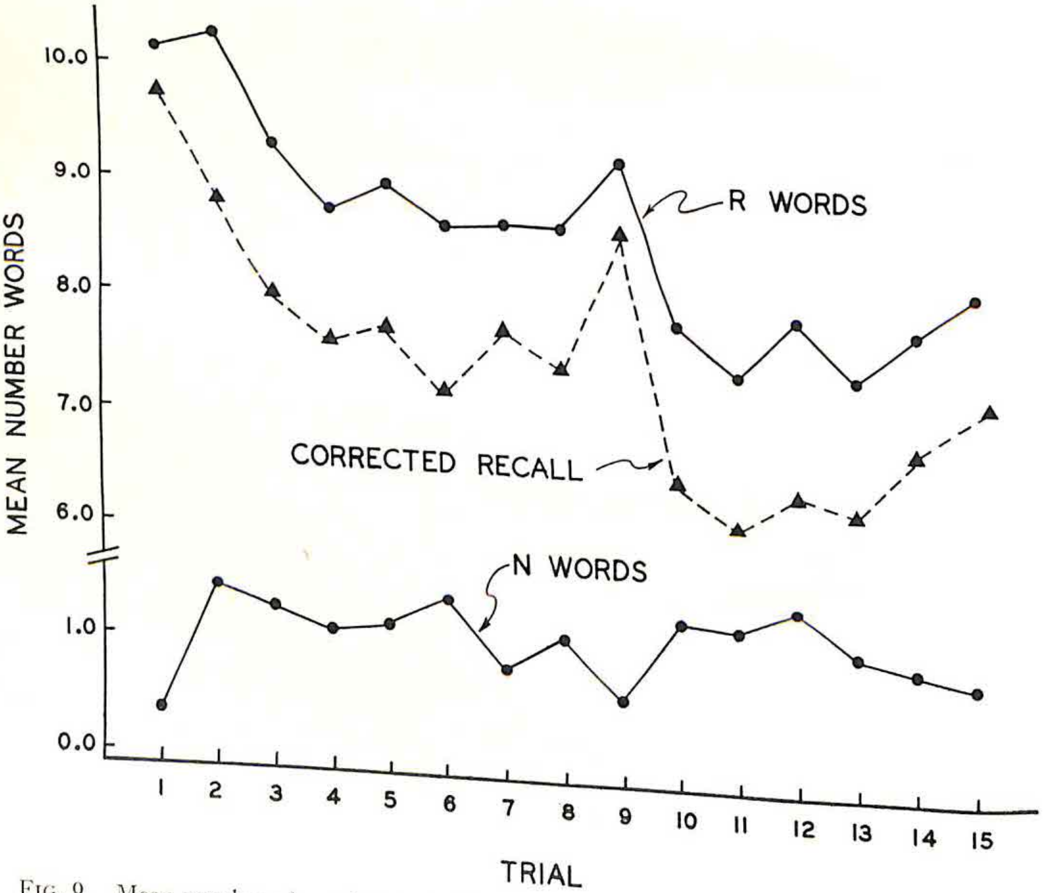


FIG. 9. Mean number of words rated "7" or "8" in Experiment IV as a function of trials.

profitably conceive of recognition and list discrimination experiments in paired-associate terms in which each word serves as the nominal "stimulus" and the List *n* marker as the nominal response. This functional identification leads to implications regarding negative transfer and retroactive and proactive inhibitions. This experiment has clearly shown negative transfer may be obtained in a list discrimination task. In G. H. Bower and J. R. Anderson (in preparation), further data will be reported to indicate that retroactive and proactive inhibition can also be obtained in a similar design.

GENERAL DISCUSSION

We will mention several conceptual puzzles regarding recognition memory which may have a viable solution within the framework of our theory. It should be stressed, however, that the theory was not constructed to account for these phenomena; they are rather in the form of afterthoughts. They will be given a more complete analysis in a forthcoming paper of G. H. Bower and J. R. Anderson.

Part–Whole Negative Transfer

The notion we have developed of negative transfer in the tagging of items in multiple lists may help to explain the puzzling phenomenon of negative transfer in part-to-whole or whole-to-part studies of free recall (see Tulving, 1966; Tulving & Osler, 1967). A subject pretrained with part of a free recall list will subsequently learn the whole list more slowly than a control subject pretrained on an irrelevant list before receiving the whole list. The difficulty is largely localized in the very poor improvement in recall of old (part-list) items (see Bower & Lesgold, 1969). This outcome would be predicted if there were negative transfer in associating a List 2 marker to an item previously associated to a number of List 1 markers, and if whole-list recall were monitored and edited for a List 2 tag associated to the candidate items before they were overtly recalled. Thus, part-list items previously associated to a List 1 tag would acquire List 2 tags more slowly and would thus be frequently edited out from free recall. This outcome hinges critically on the experimental subject not being aware that all part-list items are contained in the whole list. If he were to be informed of this fact, then there would be no list discrimination problem, and the monitor would recall any candidate item retrieved having either a List 1 or a List 2 tag associated to it. Thus, informed subjects should give only positive part-to-whole transfer. This is indeed the case, as has been found by E. Tulving (personal communication, 1971).

According to this analysis, the negative transfer in part–whole experiments is occurring in the recognition phase of free recall and not the retrieval phase. In fact, the theory expects negative transfer if the subject went from study of a whole list to more study on the selfsame whole list—provided he was led to believe that the first and second lists contained some different items. This result is precisely what has been found by R. M. Schwartz and M. S. Humphreys.[3]

Associatively Related False Alarms

The theoretical location of $f_u(x)$ relative to $f_m(x)$ in Figure 2 may provide a means for rationalizing the effects on recognition memory of similar or associatively related distractors. It is well known that such related distractors elicit more false-positive recognition judgments than do unrelated distractors. We may conceive of this as mediated recognition. Although the test item may not be directly marked, it may call to mind an associated list word which was marked, and on the basis of that evidence the subject may infer that the test item was on the list.

This indirect evidence, from mediated recognition, is available not only for related distractors but also for unmarked list items. Recall that during list study, it is presumed that the subject is searching out and marking associative pathways linking list items.

---

[3] R. M. Schwartz and M. S. Humphreys. List differentiation in part–whole free recall. Unpublished manuscript, 1971.

Hence, a list item may have several associative paths marked to other list items without itself being marked. Because mediated retrieval of a list marker can occur for either unmarked list words or associatively related distractors, it is of only partial reliability as evidence for list membership of the test item.

What the subject does with such indirect evidence will probably depend on the testing situation. For example, in multiple-choice tests, the subject would doubtless choose a directly marked test word in preference to a mediately marked test word; but he would also choose the latter over distractors having neither direct nor mediate list markings. For single stimulus or yes–no tests, allowing for mediated retrieval of list markers is equivalent mathematically to adding a constant C to the "List i evidence" scores for all items that have marked associations to items that can be independently identified as list members. This constant would be added to marked list items, unmarked list items, and associatively related distractors.

The effect of this allowance on performance depends jointly on the nature of the old study items and the new distractor items. Let us consider the base control condition to be recognition performance on a list of unrelated words tested against essentially unrelated distractors. Compared to that control, recognition would be higher for a list of highly interassociated items (e.g., members of one taxonomic category) when tested against unrelated distractors. This corresponds to what Kintsch (1968b) called "category recognition" and it is implemented in the model by increasing μ_i, the mean of the $f_i(x)$ distribution, by a constant C without at the same time shifting the distribution $f_u(x)$ for the unrelated distractors.

On the other hand, consider the case when all the distractors are in the same category (or categories) as the study items; then $f_i(x)$ and $f_u(x)$ would both be shifted up the scale by C. This is because in either case an unmarked test word of the category is likely to elicit highly associated words of that category which are marked. The result, then, would be no net change in d' or the overall recognition performance, as Kintsch (1968b) reported. In this sense, associative or categorical "organization" of the study list will affect recall but not recognition against semantically related distractors.

If the recognition test series contains a mixture of semantically related and unrelated items, then the subject is essentially dealing with three distributions having respective means of $\mu_1 + C$ for old items, $\mu_u + C$ for related distractors, and μ_u for unrelated distractors. Since the decision model supposes that the subject selects a single, item-independent criterion for making his yes–no recognition judgments, the expected outcome is a higher false-alarm probability for related than for unrelated distractors.

Returning to the earlier discussion of differentiation from unrelated distractors, the importance of list-mediated recognition will depend on the average probability of retrieving indirect list evidence for an unmarked list item during testing. That probability will depend in turn on the average degree of interitem associations established by the list-studying conditions. For example, if free classification of a set of items into more tightly packed categories improves free recall (Mandler, 1967), then the number of categories should also have a significant (though smaller) effect on list recognition against unrelated distractors. Mandler, Pearlstone, and Koopmans (1969) have reported such results; following free classification, later recognition performance (mainly hit rate) improved linearly with the number of categories into which the subjects had sorted the items.

Mandler et al. proposed a "retrieval check" hypothesis to explain their correlation between recognition and the number of categories. That hypothesis assumes that an item not recognized immediately may be recognized indirectly because it is recallable from one or another retrieval cue for the list. Our account is similar: A test item not associated directly to a list marker may be recognized because it elicits associates which do have associated list markers. Mandler et al. would seem to emphasize the associations *to* the test item *from* other list items or retrieval cues, whereas our "mediated recognition" hypothesis emphasizes the reverse associations. The two views would appear difficult to distinguish in practice.

Distinguishing Implicit Responses

An earlier criticism of familiarity theories of recognition is that they provide no basis for editing out implicit responses that subjects make to the list items, since at the time of test the two sets of items would be equated in terms of their frequency and recency. The present approach could handle such matters by supposing that the subject has control over whether or not he will try to associate an item to a list marker. Items presented by the experimenter during that temporal block denoted as "List i" activate processes which associate the items to List i markers. But implicit responses are discriminated at the time and do not activate any processes designed to associate them to List i markers. An alternative and more satisfying mechanism which achieves the same end assumes that the subject associates with the item, as one of the prevailing contextual elements, its source (the experimenter or the subject). On this latter view, the subject might be able to recall an item that helps mediate recall of other list items and have available the further information that (*a*) this word occurred in the context of List i, and (*b*) it was an implicit associate of the subject, not provided by the experimenter.

Frequency Estimates

Although the marker model was developed for handling list differentiation, it would appear also to give a usable account of frequency estimates, that is, judgments from memory of how many times a test item appeared in an earlier context. A prototypical experiment is one by Underwood, Zimmerman, and Freund (1971) in which the subject studied a long list composed of words that individually appeared either one, two, four, or six times scattered throughout the input list. During a later test series, they judged from memory how frequently a given word had appeared. This is obviously an extended recognition memory experiment, since a "zero frequency" judgment is equivalent to nonrecognition of the item, whereas a nonzero frequency judgment corresponds to recognition of the item.

Our model can keep track of item frequency by counting the different list markers associated to the item. Each time an item is presented in the context of List i, there is probability α that it becomes associated with a List i marker. Let $M_{i,n}$ denote the number of List i markers associated to an item presented n times in List i. Then $M_{i,n}$ has the binomial distribution given by

$$Pr\{M_{i,n} = x\} = \binom{n}{x} \alpha^{x}(1\text{-}\alpha)^{n-x} \quad [1]$$

The mean of $M_{i,n}$ is $n\alpha$ and the variance is $n\alpha(1 - \alpha)$, both increasing linearly with the number of presentations of an item.

A plausible hypothesis is that the subject judges the frequency of occurrence of a test item in List i by some transformation of $M_{i,n}$, the number of test markers associated to that item. The simplest mapping from $M_{i,n}$ to a frequency judgment $F_{i,n}$ is a linear transformation, namely

$$F_{i,n} = a + b\, M_{i,n}. \quad [2]$$

This relation would predict a linear function relating the mean frequency estimate to the number of presentations, a result reported by Underwood et al. (1971). Furthermore, the variance of the frequency estimates should increase linearly with the number of presentations; this prediction is approximately borne out in the Underwood et al. data.

This model may be elaborated to predict an effect of forgetting on frequency estimates. By whatever mechanism one adopts, forgetting surely produces a loss of discrimination between items presented varying numbers of times. A simple realization of this in the mathematical model is to assume that each association, established during study between an item and a list marker, has a probability $1 - f(t)$ of being retained over an interval of duration t. Therefore, as time increases, the distance between $M_{i,0}$ and $M_{i,n}$ would shrink. This could appear in the data as a decrease in the slope of the line relating mean frequency judgments to actual frequency. This prediction is upheld by the Underwood et al. data plotting the average judged versus the actual frequencies at retention intervals of a few seconds, one day, or seven days after study of the list.

REFERENCES

ABELSON, R. P., & TUKEY, J. W. Efficient utilization of non-numerical information in quantitative analysis: General theory and the case of simple order. *Annals of Mathematical Statistics*, 1963, **34**, 1347–1369.

ANDERSON, J. R. FRAN: A simulation model of free recall. In G. H. Bower (Ed.), *The psychology of learning and motivation*. Vol. 5. New York: Academic Press, 1972, in press.

BAHRICK, H. P. The ebb of retention. *Psychological Review*, 1965, **72**, 60–73.

BERNBACH, H. A. Decision processes in memory. *Psychological Review*, 1967, **74**, 462–480.

BERNBACH, H. A. Replication processes in human memory and learning. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. 3. New York: Academic Press, 1969.

BOWER, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *Psychology of learning and motivation*. Vol. 1. New York: Academic Press, 1967.

BOWER, G. H. A selective review of organizational factors in memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972, in press. (a)

BOWER, G. H. Stimulus-sampling theory of encoding variability. In E. Martin & A. Melton (Eds.), *Coding theory in learning and memory*. New York: Academic Press, 1972, in press. (b)

BOWER, G. H., & CLARK, M. C. Narrative stories as mediators for serial learning. *Psychonomic Science*, 1969, **14**, 181–182.

BOWER, G. H., & LESGOLD, A. M. Organization as a determinant of part-to-whole transfer in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 501–506.

COFER, C. N. Does conceptual organization influence the amount retained in immediate free recall? In B. Kleinmuntz (Ed.), *Concepts and the structure of memory*. New York: Wiley, 1967.

DORNBUSH, R. L., & WINNICH, W. A. Short-term intentional and incidental learning. *Journal of Experimental Psychology*, 1964, **68**, 58–63.

EAGLE, M., & LEITER, E. Recall and recognition in intentional and incidental learning. *Journal of Experimental Psychology*, 1964, **68**, 105–111.

EGAN, J. P. *Recognition memory and the operating characteristic*. (Tech. Note AFCRC-TN-58-51) Bloomington: Indiana University, Hearing and Communication Laboratory, 1958.

EHRLICH, S. Structuration and destructuration of responses in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 282–286.

ESTES, W. K. The statistical approach to learning theory. In S. Koch (Ed.) *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959.

FREUND, R. D., LOFTUS, G. R., & ATKINSON, R. C. Applications of multiprocess models for memory to continuous recognition tasks. *Journal of Mathematical Psychology*, 1969, **6**, 576–594.

GREEN, D. M., & MOSES, F. L. On the equivalence of two recognition measures of short-term memory. *Psychological Bulletin*, 1966, **66**, 228–234.

HALL, J. F. Learning as a function of word frequency. *American Journal of Psychology*, 1954, **67**, 138–140.

HINRICHS, J. V. A two-process memory-strength theory for judgment of recency. *Psychological Review*, 1970, **77**, 223–233.

HINTZMAN, D. L., & BLOCK, R. A. Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 1971, **88**, 297–306.

HINTZMAN, D. L., & WATERS, R. M. Interlist and retention intervals in list discrimination. *Psychonomic Science*, 1969, **17**, 357–358.

HINTZMAN, D. L., & WATERS, R. M. Recency and frequency as factors in list discrimination. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 218–221.

HULL, C. L., HOVLAND, C. J., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. *Mathematica-deductive theory of rate learning*. New Haven: Yale University Press, 1940.

JAMES, W. *Principles of psychology*. Vol. 1. New York: Holt, 1890.

KEPPEL, G. Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

KINTSCH, W. Recognition learning as a function of the retention interval and changes in the retention interval. *Journal of Mathematical Psychology*, 1966, **3**, 413–433.

KINTSCH, W. An experimental analysis of single stimulus tests and multiple-choice tests of recognition memory. *Journal of Experimental Psychology*, 1968, **76**, 1–6. (a)

KINTSCH, W. Recognition and free recall of organized lists. *Journal of Experimental Psychology*, 1968, **78**, 481–487. (b)

KINTSCH, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

Kintsch, W., & Carlson, W. J. Changes in the memory operating characteristic during recognition learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 891–896.

Kjeldergaard, P. M. Transfer and mediation in verbal learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Lockhart, R. S., & Murdock, B. B., Jr. Memory and the theory of signal detection. *Psychological Bulletin*, 1970, **74**, 100–109.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. 1. New York: Academic Press, 1967.

Mandler, G., Pearlstone, A., & Koopmans, H. S. Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 410–423.

McDougall, R. Recognition and recall. *Journal of Philosophical and Scientific Methods*, 1904, **1**, 229–233.

McGovern, J. B. Extinction of associations in four transfer paradigms, *Psychological Monographs*, 1964, **78** (16, Whole No. 593).

Müller, G. E. Zur Analyse der Gedächtnistätigkeit und des Vorstellungsverlaufes. *Zeitschrift fur Psychologie*, Erganzungsband 8, 1913.

Murdock, B. B., Jr. The immediate retention of unrelated words. *Journal of Experimental Psychology*, 1960, **60**, 222–234.

Murdock, B. B., Jr. Signal detection and short-term memory. *Journal of Experimental Psychology*, 1965, **70**, 443–447.

Norman, D. A., & Rumelhart, D. E. A system for perception and memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

Olson, G. M. Learning and retention in a continuous recognition task. *Journal of Experimental Psychology*, 1969, **81**, 381–384.

Parks, T. E. Signal-detectability theory of recognition-memory performance. *Psychological Review*, 1966, **73**, 44–58.

Postman, L. One-trial learning. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963.

Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and method of testing. *Journal of Experimental Psychology*, 1955, **49**, 1–10.

Quillian, M. R. The teachable language comprehender. *Communications of the Association for Computing Machinery*, 1969, **12**, 459–476.

Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972, in press.

Schwartz, F., & Rouse, R. D. The activation and recovery of associations. *Psychological Issues*, 1961, **3** (Whole No. 9).

Shepard, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 156–163.

Shiffrin, R. M. Forgetting: Trace erosion or retrieval failure? *Science*, 1970, **168**, 1601–1603. (a)

Shiffrin, R. M. Memory search. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970. (b)

Tulving, E. Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 193–197.

Tulving, E., & Osler, S. Transfer effects in whole-part transfer. *Canadian Journal of Psychology*, 1967, **21**, 253–262.

Underwood, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychological Monographs*, 1945, **59** (3, Whole No. 273).

Underwood, B. J., Zimmerman, J., & Freund, J. S. Retention of frequency information with observations on recognition and recall. *Journal of Experimental Psychology*, 1971, **87**, 149–162.

Wickelgren, W. A., & Norman, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, **3**, 316–347.

Winograd, E. List differentiation as a function of frequency and retention interval. *Journal of Experimental Psychology*, 1968, **76** (2, Pt. 2).

(Received June 14, 1971)

processes. As an example, any language spoken at a normal intensity is easily detected but impossible to understand without the appropriate knowledge in long-term memory. Similar to visual masking (Haber, 1969; Kahneman, 1968; Raab, 1963), the auditory masking literature also supports the distinction between detection and recognition.

Auditory Detection Masking

A number of studies have been carried out studying backward masking of auditory detection (Deatherage & Evans, 1969; Elliot, 1962a, 1962b, 1967; Homick, Elfner, & Boothe, 1969; Raab, 1963; Samoilova, 1959). For example, Elliot (1967, Experiment I) employed a backward masking paradigm to study the development of auditory frequency contours. Durations of 10 and 100 milliseconds (msec.) were used for the test tone and masking noise, respectively. The steady-state level of the masking noise was 70-decibels (db.) sound pressure level (SPL). Thresholds for the detection of the test tone were computed at backward masking intervals of 10 to 100 msec. Backward masking decreased with increases in the masking interval. Very little masking was observed when the masking noise followed the test tone by 100 msec.

Elliot's (1967) results are supported by a number of recent detection masking experiments (Deatherage & Evans, 1969; Homick et al., 1969). The results indicate that a loud masker occurring within 100 msec. of a test stimulus presentation increases the detection threshold of the test stimulus. Further, the detection threshold is inversely related to the duration of the silent interval between the test and masking stimuli. However, these studies of backward masking do not necessarily provide evidence for a preperceptual auditory store. Rather, the masking stimulus could overtake the test stimulus in the auditory pathways decreasing the signal-to-noise ratio and, therefore, decreasing detection performance. In detection masking studies, the masking stimulus is much louder than the test stimulus. The question of interest, then, is how the loudness of a stimulus relates to the time it takes to be detected. McGill (1961, 1963) has demonstrated an inverse relationship between stimulus loudness and simple reaction time. Assuming that simple reaction time to a tone contains independent detection, decision, and response components (Donders, 1969; Sternberg, 1969), stimulus loudness should only affect the detection component. Therefore, it is reasonable to account for the decision and response components and to measure relative detection time as a function of stimulus loudness.

McGill (1961) presented a subject with a 1,000-hertz (Hz.) tone of various random intensities in a simple detection task. McGill's data indicated that increasing the amplitude of the test tone from 30- to 100-db. SPL decreased the median reaction time from 216 to 120 msec. This difference, which represents differences in detection time, approximates 100 msec. which is exactly the value usually found as the effective duration of backward detection masking. Therefore, the temporal course of backward detection masking could be equal to the difference in the detection times of the test and masking stimuli. If the masking stimulus overtakes the test tone in the auditory pathways, the difference in reaction times to the test and masking stimuli should equal the temporal course of backward masking.

Most quantitative studies of backward detection masking have employed test tones and masking noise as stimuli. Simple reaction times to noise of a certain intensity may not equal that of a tone of the same intensity. Keeping this in mind, it still may be helpful to interpret the effect of backward masking in terms of the detection times given by McGill (1963). Homick et al. (1969) studied backward masking of a tonal signal as a function of the intensity of the masking noise. The results indicated that increasing the masking noise from 70-db. to 90-db. SPL increased the temporal course of backward masking. For example, the 90-db. masker produced about the same amount of masking as the 70-db. masker if the 90-db.

masker was delayed an extra 20 msec. For example, subjects could detect a given test signal if there was silent interval of 30 msec. before the 70-db. masking noise. However, with a masking noise of 90 db. subjects needed a silent interval of 50 msec. for correct detection of the test signal. Accordingly, it is possible that the 70-db. noise takes about 20 msec. longer to be detected than the 90-db. noise. McGill's relative reaction times to 70- and 90-db. tones support this. It took McGill's subjects about 20 msec. longer to make a simple response to a 70-db. tone than to a 90-db. tone. The results indicate that backward detection masking occurs because the second masking stimulus overtakes the test stimulus and decreases the signal-to-noise ratio.

The results of backward detection masking do not provide evidence for the existence of a preperceptual auditory store. The present interpretation of backward detection masking must now be reconciled with a different account of forward detection masking. Forward masking of detection has also been demonstrated in a number of different experiments (Deatherage & Evans, 1969; Elliot, 1962a, 1962b; Homick et al., 1969). Although Elliot (1962a, 1962b) found less forward masking than backward masking, Homick et al. (1969) found at least as much forward masking as backward masking. It was argued that backward detection masking occurs because the masking stimulus overtakes the test stimulus decreasing the signal-to-noise ratio. In forward masking, this cannot be the case since the louder masking stimulus is presented first and would be detected before the test stimulus. Since the masking stimulus is terminated before the test tone is presented, no masking should take place unless an auditory image of the masking stimulus outlasts its presentation.

Forward masking can be accounted for by assuming that an auditory image outlasts the masking stimulus presentation. Results that will be presented later provide direct demonstrations of an auditory image that remains after a stimulus is presented. The results also indicate that the auditory image that outlasts the stimulus does not differ qualitatively from the image during the stimulus presentation. Therefore, the auditory image of a forward masking stimulus should interfere with detection of a test stimulus in the same way as a simultaneous masking stimulus (Green & Swets, 1966). Accordingly, a forward masking stimulus interferes with detection, since the auditory image of the masking stimulus decreases the signal-to-noise ratio of the test stimulus.

Erikson and Johnson (1964) have been credited with demonstrating that preperceptual echoic memory might last as much as 10 seconds (sec.) (Bryden, 1971; Neisser, 1967). Erikson and Johnson (1964) presented subjects with a barely detectable tone while they were engaged in reading a novel. The reading light was turned off either simultaneously with the tone or from 1 to 10 sec. after the tone presentation. To keep the subjects honest, the light was also turned off at times when no tone had been presented. At the offset of the light, the subjects had to report "whether or not the tone had sounded sometime within the last 10–15 sec. [p. 30]." The results indicated that the ability to correctly report whether a tone had been presented decreased with increases in the tone–test interval. However, subjects could still reliably report a tone presentation 10 sec. after its presentation. Neisser (1967) concludes from these results that "If we assume that attention to the novel precluded any encoding the of 'beep' we can infer that recall in this procedure must have been based on echoic memory alone [p. 204]."

However, Neisser's (1967) assumption cannot be justified and probably could easily have been disproved by Erikson and Johnson's own subjects. A phenomenological report could have indicated whether subjects noticed the tone when it was presented. Furthermore, at the offset of the light did they search for an auditory image or for a recent awareness or covert identification of the tone presentation? As Massaro (1970b) points out, the recall decision might have been based on memory for the covert identification rather than on storage of the preperceptual image. Fortunately, one of the experimental conditions helps resolve the issue. Introducing a roving frequency tone sometime after the tone presentation did not decrease accuracy of report. In contrast, the masking results discussed below indicate that preperceptual store is easily disrupted by subsequent auditory inputs. Therefore, the reports of Erikson and Johnson's subjects were not based on a preperceptual image present at the time of the test.

Forward and backward detection masking studies do not provide direct evidence for the existence of a preperceptual auditory store. Furthermore, since most auditory inputs outside the laboratory are easily detected, the detection masking studies are not directly relevant to the temporal course of processing familiar auditory information. Quantifying the temporal course of masking would be relevant to the temporal course of processing auditory information only if the masking stimulus terminated processing of that information rather than prevented the detection of that information. For example, a second input in speech does not prevent detection of the first, but could terminate the perceptual processing of the first. This phenomenon would be similar to one in vision in which an eye movement while reading is sufficient to erase the earlier pattern. As in vision, the detection masking paradigm is not appropriate for studying the temporal course of auditory identification or recognition. Accordingly, a recognition masking task was devised to determine the properties of the auditory information storage system and to quantify the temporal course of processing auditory information (Massaro, 1970b, 1971).

Auditory Recognition Masking

In the recognition paradigm, the subject first learns to identify or recognize

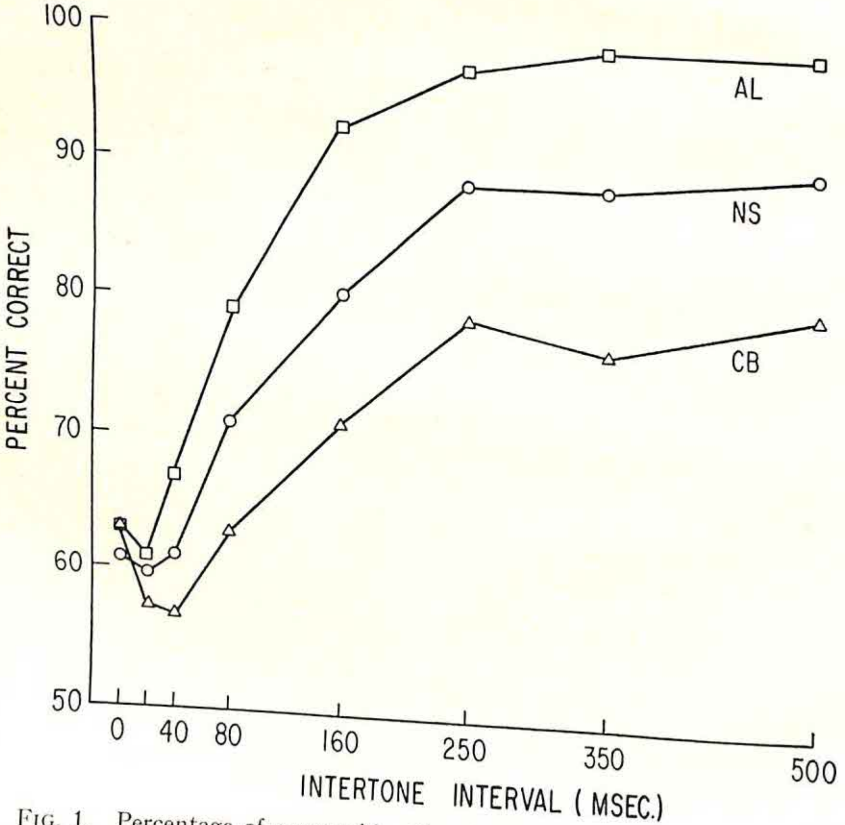


FIG. 1. Percentage of correct identifications of the test tone for subjects AL, NS, and CB as a function of the duration of the silent intertone interval.

two or more test signals. For example, the test signals could be two short tones differing in pitch. The subject's task is to identify the higher tone as high and the lower tone as low. In the backward masking paradigm, one of the test tones is presented followed by a silent interval followed by a masking tone. The test and masking tones are presented at the same loudness so that the masking stimulus will not overtake the test tone as in backward detection masking studies.

Massaro (1970b) presents an experiment that can serve as a prototype for discussion. In this study, one of two pure tones (a 20-msec. sine wave of 770 or 870 Hz.) was presented and the subjects' task was to identify the tones as low and high, respectively. The masking tone was equal to 820 Hz. All tones were presented at 81 db. The silent intertone interval lasted 0, 20, 40, 80, 160, 250, 350, or 500 msec. The masking tone lasted 500 msec.

Figure 1 shows that for each of the three subjects in the task, recognition performance improved with increases in the silent intertone interval up to 250 msec. Further increases in the silent interval beyond 250 msec. did not significantly facilitate recognition performance. These results provide information about the preperceptual auditory image of the test tone, the vulnerability of the auditory image to new inputs, and the temporal course of recognition. Given that the test tone lasted only 20 msec., some preperceptual image must have remained for perceptual processing necessary to improve recognition performance with increases in the silent intertone interval. This same result indicates that the masking tone terminated perceptual processing of the image. Since recognition performance levels off at about 250 msec., the image probably decayed within this period. These results also indicate that the subjects required at least 250 msec. of silence besides the 20 msec. of the test tone presentation for the recognition process.

Differences between Detection and Recognition Masking

A number of comparisons indicate that the backward masking observed in the recognition paradigm reflects different processes than those assumed to operate in detection masking. First, Elliot (1967) and Homick et al. (1969) have shown that the amount of backward detection masking is dependent on the frequency similarity of the test tone to the center of the masking noise. The results indicated that at all masking intervals, increasing the similarity of the narrow band masking noise to the test tone increased the total amount of masking. This result agrees with the assumption that the masking stimulus in detection masking decreases the signal-to-noise ratio. Simultaneous masking results also indicate that increasing the similarity of the test and masking stimulus decreases detection performance. In contrast, Massaro (1970b) found that the disrupting effect of the masking tone was not dependent on the frequency similarity of the masking tone to the test tone. This result agrees with the assumption that the masking tone in the recognition task terminates perceptual processing of the auditory image.

Second, presenting the masking tone contralateral to the test tone presentation does not decrease backward masking in the recognition paradigm (Massaro, 1970b). This result indicates that the preperceptual auditory image is located centrally rather than at the peripheral level. In contrast, a dichotic stimulus produces very little backward masking in the detection paradigm (Deatherage & Evans, 1969; Elliot, 1962a, 1962b). In agreement with simultaneous detection masking (Deatherage & Evans, 1969; Fletcher, 1953), a dichotic masker does not reduce the signal-to-noise ratio and detection is disrupted very little.

Third, although foward masking decreases detection performance significantly (Homick et al., 1969), Massaro (1970b) found no forward masking in the recognition paradigm. It was concluded earlier that forward masking of detection was due to the decreased signal-to-noise ratio produced by the auditory image. In contrast, forward masking does not occur in recognition, since the new image of the test tone terminates processing of the image of the masking stimulus. This occurs since both stimuli in the recognition task are equal in loudness. In contrast, the test stimulus in forward detection masking does not terminate processing of the image of the masking stimulus, since the test stimulus is much softer than the masking stimulus. The comparisons between detection and recognition masking show that detection masking is a phenomenon that is closely tied to the psychophysical processes operating in simultaneous masking studies. Recognition masking, on the other hand, is closely tied to the temporal course of perceptual processing.

The recognition masking results indicate that a short tone presentation leaves an auditory image which decays very rapidly. Further, this image can be processed for correct identification. A second new stimulus interferes with perceptual processing necessary for correct recognition. Two important questions raised by the recognition masking results need to be answered. First, what is the nature of the image that remains after presentation of a short auditory stimulus? Second, how does the masking stimulus interfere with perceptual processing of the image?

Preperceptual Auditory Images

Using a paradigm similar to one employed by Sperling (1967) and Haber and Standing (1969) for visual stimuli, Efron (1970a, 1970b, 1970c) provides an analysis of the perceived duration of tones and noise. In this paradigm, subjects judge the temporal overlap between two stimuli presented sequentially. The subjects state whether or not the onset of the second index stimulus (e.g., a light) occurred before the offset of the first auditory stimulus. The interval between the offset of the auditory stimulus and the onset of the index stimulus is adjusted until a subject perceives offset–onset simultaneity. The independent variable of interest is the duration of the first auditory stimulus. The results indicate that the minimal perception of an auditory stimulus lasts about 130 msec. Decreasing the duration of the first auditory stimulus below 130 msec. increased its perceived duration by a similar amount. For example, if the first noise burst lasted 30 msec., the subject did not perceive a temporal interval between its offset and the onset of the index stimulus until there was an interstimulus interval of 100 msec. Efron's (1970a, 1970b, 1970c) results support the assumption of preperceptual auditory images. If a short tone produces an auditory image for perceptual processing, the subject should be able to estimate its duration. Finally, analogous to visual information processing (Haber & Standing, 1969; Sperling, 1960, 1967), the auditory afterimage does not differ qualitatively from the image during the stimulus presentation.

Plomp (1964) studied the rate of decay of auditory sensation in a two-alternative task. He presented a 200-msec. pulse of noise followed by another noise burst either immediately or after a silent interval. Given that a noise burst produces an afterimage that decays over time, the subject should not notice a blank interval before a second softer burst of noise. Further, the length of the blank interval that goes unnoticed should increase with intensity differences between the two noise bursts. Plomp showed that a time interval of 2.6 msec. is noticed 75% of the time between two noise bursts of 65 db. In contrast, when the second noise burst is reduced to 15 db., the blank interval must be increased to 78 msec. to be noticed 75% of the time. This result indicates that the auditory image of the first noise burst decayed to 23% of its original value in 78 msec. The two-alternative task employed by Plomp controls for decision biases, whereas Efron's (1970a, 1970b, 1970c) paradigm does not. Therefore, it seems necessary to employ Plomp's task while varying the duration of the first noise burst to provide another test of Efron's results that the duration of the afterimage is inversely related to the duration of the test stimulus.

Some Russian studies have also provided evidence for an auditory image that remains after a short tone is presented (Gol'dburt, 1961). Gol'dburt has shown that a second tone can shorten the perceived duration of an earlier short tone. Perceived duration of the first tone increases with increases in the silent interval between the two tones. This result indicates that the auditory image extends the apparent presentation time of a short tone. A second tone can interfere with the auditory image and decrease the perceived duration of the first tone. Gol'dburt (1961) has also shown that the effect of the second tone decreases with increases in the duration of the first tone. This indicates that analogous to the visual image (Haber & Standing, 1969), the duration of an auditory image is inversely related to the duration of the stimulus producing the image. If perceptual processing usually takes about 250 msec., there would be no need for an image to remain if the presentation time of the stimulus exceeds this value. Processing the information in the stimulus seems to be sufficient to eliminate any afterimage of the stimulus presentation.

Von Békésy (1971), simulating backward inhibition in concert halls, measured the perception of a sound that was followed by a second sound in a different location. He presented a 1,000-Hz. tone for 35 msec. over a single speaker followed by a 1,500-Hz. tone over a ring of speakers surrounding the single speaker. If the second tone is presented 60 msec. after the onset of the first tone, the second tone reduces the perceived duration and loudness of the first tone. Von Békésy states,

> But if we assume that every stimulus starts a process in the brain which lasts perhaps 200 milliseconds, we can make backward inhibition acceptable if we further suppose that this process can be inhibited at any moment during the 200-millisecond interval by the onset of the second stimulus [p. 530].

The studies of Efron (1970a, 1970b, 1970c), Plomp (1964), Gol'dburt (1961), and von Békésy (1971) indicate that the auditory image remaining after a short stimulus does not differ significantly from the image present during the stimulus presentation. A short auditory stimulus usually requires perceptual processing that outlasts the life of the stimulus. The auditory image remaining after the stimulus presentation contains the necessary information for perceptual processing. Since the afterimage does not differ from the image present during the stimulus, the subject mistakes the afterimage for the image that is present during the stimulus presentation. Accordingly, the subject overestimates the duration of a short stimulus. With a longer stimulus, the subject has time to process the information available during the stimulus presentation and does not overestimate the duration of the stimulus.

In a different approach to measuring the duration of preperceptual auditory images, Guttman and Julesz (1963) repeated an identical section of wide-band noise. The authors varied the duration of the section and measured this effect on periodicity perception. It is difficult to judge whether this paradigm is appropriate for determining the duration of preperceptual auditory images. However, the results and an interpretation will be presented for completeness. If a section of white noise is repeated above 19 cycles per second (cps), the periodicity is heard as pitch. Between 4 and 19 cps, motorboating is heard. In these two cases, the stimulus sounds smooth or homogeneous within a period. Although listeners can hear a whooshing in the range of 1–4 cps, an intraperiod roughness is heard. Which of these estimates qualify for estimating the duration of auditory images? If subjects perceive a change within a period, they have already synthesized some of the preperceptual image. Therefore, more than the preperceptual image is responsible for the perception of whooshing at 1–4 cps. The motorboating at 4–19 cps sounds homogeneous within a period. The duration of preperceptual images should be measured with the restriction that the noise section sound homogeneous within a period since synthesis has not taken place. Employing this criterion, the results indicate 250 msec. as the maximal duration for preperceptual auditory images.

Massaro (1972) has provided another demonstration that the auditory image does not differ qualitatively from the image present during the stimulus presentation. In the recognition masking paradigm, the question of interest was which variable predicts tone recognition best: test stimulus duration or processing time. Although increasing the duration of a tone had a slight facilitatory effect on identification performance, processing time was most critical for accurate identification. That is, with processing time held constant, increasing the duration of the test tone had very little effect on identification performance. On the other hand, increasing processing time improved identification performance independent of the duration of the test tone. In terms of a perceptual processing model (Massaro, 1970a), the rate of processing the information in a test stimulus presentation did not differ as a function of following a short tone presentation by a silent period or by simply leaving the test tone on for the processing interval.

Kahn and Massaro[3] have shown that the afterimage of the test tone must have acoustic properties similar to the image of the test tone presentation. They reasoned that it was logically possible that the masking tone simply distracted the subject's attention and reduced recognition performance. If masking was due to switching of attention, the processing of a nonauditory stimulus should also function as a masking stimulus in the tone recognition task. To test this, light and tone masks were employed in the tone identification task. The subjects were also required to identify the duration (long or short) of the masking stimulus to make certain that they would attend to the masking stimulus when it was presented.

Following the test tone with a masking tone replicated earlier studies indicating that identification performance improved with increases in the silent intertone interval. On the other hand, when the test tone was followed by a masking light, test tone identification did not depend on the interstimulus interval and was at an asymptotic level. The accurate identification of the duration of the masking stimuli indicated that the subjects processed the masking stimuli. Therefore, these results support the conclusion that the auditory image is stored with the acoustic features of the test stimulus presentation and these features are not lost by a simple switch in attention. A second auditory input interferes with recognition by interfering directly with the auditory information available from the earlier test stimulus presentation.

The results support the first assumption that an auditory input produces a preperceptual auditory image that contains the information in the auditory stimulus. The image persists beyond the stimulus presentation and preserves its acoustic information. The second assumption is that the recognition process entails a readout of the information in the preperceptual auditory image. The recognition process takes time and is referred to as the temporal course of perceptual processing. The following discussion provides an analysis of the time required for perceptual processing.

Perceptual Processing Time

An experimental paradigm has been developed to establish the minimal duration necessary to perceive a repeated sequence of auditory items. Warren, Obusek, Farmer, and Warren (1969) have shown that subjects need at least 300 msec. of each of four nonverbal stimuli for accurate perception of their temporal order when the items are repeated in a loop. The experimental task involves identification of a repeated sequence of four successive sounds (e.g., high tone, buzz, low tone, and hiss). Although inexperienced subjects need much longer, experienced subjects cannot accurately identify the correct temporal order unless each stimulus lasts at least 300 msec.

The fact that four repeated digits (Warren & Warren, 1970) or four repeated vowels (Thomas, Hill, Carrol, & Bienvenido, 1970) can be recognized at much shorter durations (125–200 msec.) suggests that there may be some qualitative differences in perception of speech and nonspeech material. On the other hand, as Thomas et al. point out, the time to reach a decision for familiar sounds (speech) would be expected to be much less than the time required for unfamiliar nonspeech stimuli. A direct comparison of the processing times required for speechlike and nonspeechlike material while controlling for familiarity (i.e., discriminability) has not been reported. It seems unlikely that the temporal course of pattern recognition would be qualitatively different for speech and nonspeech material if the familiarity and distinctiveness of the material are taken into account.

---

[3] B. J. Kahn and D. W. Massaro. Backward recognition masking: Interference or shift of attention. Unpublished manuscript, 1971.

A measure of processing time in speech might be found in the duration of vowels in normal speech which are in the range of 150 to 350 mesc. (Fletcher, 1953; House, 1961). Since the pattern of sound pressure changes in a steady state vowel repeats at the speaker's fundamental frequency (Fletcher, 1953), the extended duration of the vowel might be needed for processing the information available in the vowel presentation. Vowels at much shorter durations can be identified if followed by a silent retroactive interval (Gray, 1942). However, if processing is interfered with by following the short test vowel presentation with other vowels, the test vowel should not be identified. This result would provide evidence that the redundancy of the vowel in normal speech allows time for processing, since the extended duration of the vowel protects it from later speech until processing has been completed.

Massaro[4] has also employed short vowel stimuli in the recognition masking paradigm described earlier. The spoken vowels were recorded at the same fundamental frequency and amplitude. A steady-state segment of each vowel was stored digitally employing a computer-controlled analog-to-digital converter. During the experiment, the vowel segment was played back using a digital-to-analog converter. In the recognition masking task, 20-msec. segments of the vowels /i/ as in "heat" and /I/ as in "hit" were employed as test items. The masking stimulus was a 270-msec. nonsense vowel made up of two alternating vowel segments. These two segments were taken from the vowels /a/ as in "hat" and /U/ as in "put" and lasted 45 msec. each. The subjects had three days of practice identifying the test vowels in this task before the present experiment.

The results of the experiment are shown in Figure 2. For each subject, identification performance increased with increases in the silent intervowel interval. Therefore, the results indicate that processing

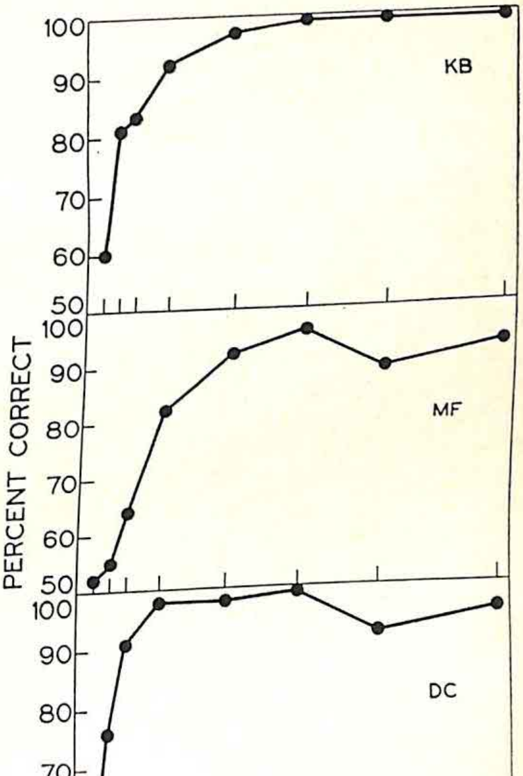


FIG. 2. Percentage of correct identifications of the test vowel for subjects KB, MF, and DC as a function of the duration of the silent intervowel interval.

---

[4] D. W. Massaro. Perceptual processing time of short vowels. Unpublished manuscript, 1971.

time is also critical for vowel identification. Performance improved much faster with increases in the silent interval in this task than in the tone identification task (cf. Figure 1). The effective masking interval for pure tones was almost twice as long as the masking interval for vowels. This comparison indicates that "i" and "I" are more distinctive than tones of 870 and 770 Hz. Therefore, the vowels could be identified much faster than the tones and were less sensitive to masking at longer silent intervals.

It should be possible to increase the effective masking interval of vowels by increasing the number of test alternatives. With only two vowels subjects can probably make their decision very rapidly, since few features of the vowels need be processed for accurate identification. With a larger number of alternatives, however, subjects would have to process more features before reaching a decision. Accordingly, increasing the number of alternatives should increase the temporal course of backward masking.

The estimates of perceptual processing time from a number of different tasks agree remarkably. The recognition masking studies indicate that perceptual processing can last between 120 and 250 msec. Identifying a sequence of four auditory stimuli requires anywhere from 125 to 300 msec. per stimulus depending on the nature of the material. Although these estimates were obtained in fairly simple experimental situations, they should not be unrelated to the temporal processing of complex auditory information. Therefore, these results will be important in the discussion of processing continuous speech.

The estimates of perceptual processing time are in the same range as the estimates of the duration of preperceptual auditory images. The first 100 to 250 msec. of a stimulus presentation is critical for the recognition process. An auditory stimulus presented for shorter periods produces an auditory image. Accordingly, perceptual processing of a short stimulus continues during the retroactive silent interval. In the recognition masking studies, the short tone or vowel presentation functions as a perceptual unit and recognition corresponds to a readout of information in the unit. If an auditory stimulus changes rapidly over time (e.g., during a consonant-vowel transition), perceptual processing must continue after presentation of the perceptual unit. Therefore, it is necessary to specify what stimulus variables determine the formation of a perceptual unit of information.

Perceptual Units

According to the present theory, a perceptual unit cannot exceed preperceptual auditory memory and must be followed by a steady-state or silent period so that perceptual processing can take place. Perceptual units must also be identified with respect to the temporal changes in sound pressure of the auditory signal. In recognition masking, the test stimulus functioned as the perceptual unit. It is interesting that the masking stimulus could not be integrated with the test stimulus, forming a perceptual unit, and therefore reduce the masking effect. Some integration of the test and masking tones may have occurred when the masking tone followed the test tone immediately. This would account for the initial rise in performance at the zero silent interval in Figure 1.

If the masking tone could be integrated with the test tone, it would not necessarily decrease identification performance. The stimulus situation can be constructed so that integration of two stimuli will be possible as can be seen in Hirsh's (1959) study. In Hirsh's task, the subject is required to judge which of two sounds came first. The onset difference between the two sounds is varied while keeping their offsets simultaneous. Employing two pure tones that differ in frequency, Hirsh showed that the first tone can be identified 75% of the time if it precedes the second by 17 msec. At first glance, this paradigm seems remarkably similar to the auditory recognition masking task discussed above. In this case, Hirsh's results seem discrepant with Massaro's (1970b) results that indicated that it takes about 1/4 sec. to process the pitch of a test tone. However, it appears that subjects are able to integrate two tones that overlap in time, thus reducing the masking effect of the second tone. Accordingly, the perceptual unit in Hirsh's task was the quality of the pair of tones (Broadbent & Ladefoged, 1959). With high and low tones, the subject learns the two integrated stimuli: high-low and low-high. By simply identifying the integrated stimulus, the subject can indicate which tone came first. On the other hand, if a silent interval occurs between the two tones as in Massaro's (1970b) task, integration cannot take place and the single test tone functions as the perceptual unit. In this case, the second tone terminates perceptual processing of the test tone.

The overall duration of the tones in Hirsh's experiments varied between 440 and 560 msec. There is also evidence that subjects can perceive temporal order at even smaller intervals if the overall duration of the tones is shortened (Patterson & Green, 1970). By shortening the overall duration of Hirsh's low and high tones to 10 msec., Patterson and Green showed that subjects can identify a temporal onset difference of 1.5 msec. This temporal interval is now similar to the empty interval necessary for separation between two brief sounds for subjects to report there are two sounds instead of one (Miller & Taylor, 1948). Hirsh (1959) assumed that the perceptual process involved in identifying whether one or two sounds have occurred is at a lower level than the process involved in the temporal order task. Hirsh felt that since 20 msec. is needed for judging temporal order, more of the perceptual system is involved than the ear itself. However, we can reject Hirsh's distinction between these two tasks since Patterson and Green's results indicate that very short time differences can also be recognized in the temporal order task. Two milliseconds cannot be a measure of central processing time, since this would require that perception be essentially immediate.

Perceptual processing is the identification of acoustic information that is probably stored sequentially. The visual image prolongs the stimulus in its correct spatial pattern (Sperling, 1960). If the auditory image does something analogous to the visual image, it must preserve the stimulus in its original sequential pattern for perceptual processing. The time to process the auditory input in the temporal order task should not be in the range of the onset differences that can be discriminated, but rather should reflect the processing time measured in the pitch identification task (Massaro, 1970b), To measure central processing time, it is necessary to stop perceptual processing at varying times after stimulus presentation. Accordingly, to disrupt performance in the temporal order task, two brief tones differing in onset times could be followed with a third tone after a short silent period. From Massaro's results, the third tone should mask the identification of temporal order. Furthermore, the identification of temporal order might be vulnerable for about 250 msec. after the two tones are presented.

Studdert-Kennedy, Shankweiler, and Schulman (1970) studied the identification of two consonant-vowel syllables as a function of their differences in onset time. The two syllables only differed with respect to the initial consonant, since they were chosen from the set of six syllables /ba,da,ga,pa,ta,ka/. On each trial, the two syllables were presented for 250 msec. to the same ear or to different ears. For both monaural and dichotic presentation, subjects identified a given syllable about 65% of the time if the two syllables were presented simultaneously. Under monaural presentation, increasing the differences in onset time improved identification of the first syllable and disrupted identification of the second syllable. In contrast, increasing the differences in onset time improved identification of both syllables under dichotic presentation.

These results indicate different processes operating in the monaural and dichotic presentations. Monaural presentation of the syllables produces a stimulus situation similar to the paradigm employed by Hirsh (1959). Supporting this, the monaural data agree with Hirsh's results. Subjects were able to identify the first syllable 75%

of the time if it led the second by about 17 msec. If the first syllable led by 50 msec. or more, its identification was perfect. Other results provide evidence for simultaneous masking effects with monaural presentation. Identification of the second syllable decreased with increases in onset differences between the two syllables. Identification of the second syllable is near chance if it follows the first by 50 msec. or more. Studdert-Kennedy et al. (1970) point out that with onset differences of 50 msec. or more, the stimulus sounds like a single syllable with a superimposed click.

Identification of the syllables with dichotic presentation is a different matter. The second syllable is identified about 85% of the time if it follows the first by 50 msec. or more. The fact that the second syllable cannot be identified at these lag times in monaural presentation indicates that dichotic presentation preserves the integrity of the syllable at each ear. Accordingly, since dichotic presentation separates the two syllables into two perceptual units, this task is more similar to Massaro's (1970b) recognition paradigm. Therefore, presentation of the second syllable should interfere with identification of the first. This follows from the assumption that a second input that cannot be integrated with the first terminates perceptual processing of the first. The results support this assumption. Identification of the first syllable was lowest if it led the second by 20 msec. and improved with increases in lead time. Identification of the first syllable was still only 80% correct when it led the second syllable by 120 msec. Further increases in lead time would have increased identification of the first syllable. Extrapolating the results indicates that the optimal processing time for the first syllable would be close to the 250 msec. found by Massaro (1970b) for pure tones. Therefore, the results of Studdert-Kennedy et al. (1970) support the importance of perceptual units and perceptual processing time in auditory perception.

Creel, Boomsliter, and Powers (1970) present clinical evidence that indicates that the perception of tones or noise is temporally categorical. Patients with a deficient blood supply to the brain stem need as much as 200 to 400 msec. of a 1,000-Hz. tone to perceive it as tone rather than noise. A normal listener needs about 10 msec. of the same burst to achieve the tonal sensation (Stevens & Davis, 1938). A given patient hears a sine wave presented for 300 msec. as noise and the same sine wave presented for 400 msec. as tone. As the authors point out, given a tone of 400 msec., one might expect that this patient would hear 300 msec. of noise followed by a short tonal sensation. Since the perception is either noise or tone, the signal is perceived as a Gestalt and is therefore temporally categorical. This result provides convincing evidence that perception is not immediate and that auditory perception is the result of processing preperceptual units of information.

Even though 10 msec. is a sufficient duration for a tone to be heard as tone for normal observers, the results discussed earlier indicate that it takes much longer to process the information necessary for the tonal sensation. Patients with an impaired blood supply to the brain may need more processing time than normals for a tonal sensation. Since the image of a short tone decays rapidly, its information may not last long enough for the patient's perceptual processing necessary for a tonal sensation. In terms of analysis by synthesis (Neisser, 1967), the information may not be present long enough for synthesis of a tonal sensation. Extending the duration of the tone, of course, increases the life of the information necessary for processing and a pure tone is perceived.

Processing Continuous Speech

Studies of interrupted and alternated speech also provide evidence concerning perceptual units. In interrupted speech, half of the speech is eliminated by replacing segments of the speech signal with silence. Therefore, speech and silence are alternated at a given rate. Miller and Licklider (1950) and Dirks and Bower (1970) have shown that intelligiblilty of speech was not lowered significantly by eliminating half of it if the empty intervals occur every 50 or 5 msec. However, if the speech was replaced with silence every other 500 msec., about half of the words were recognized. These results indicate that the durations of the perceptual units critical for the monosyllabic words employed are between 50 and 500 msec. Furthermore, the redundancy of the perceptual unit makes recognition possible even though alternate portions of it are replaced with silence.

Huggins (1964) studied the perception of speech passages taken from scientific essays for the layman. Under conditions of interrupted speech, he found that speech intelligibility was lowest at rates of 1.5 to 5 interruptions per second. Therefore, the poorest speech recognition is produced by replacing the speech signal with silent periods that last between 100 and 330 msec. Since the durations of syllables are within this range, the results implicate the syllable as the perceptual unit for speech. If the syllable functioned as the perceptual unit for speech recognition, its complete removal should and does lead to the poorest speech recognition.

Alternating continuous speech from ear to ear rather than removing segments of the speech signal also provides information about the perceptual unit of speech perception. Alternating speech from ear to ear about three times per second disrupts speech perception (Cherry, 1953; Cherry & Taylor, 1954). In contrast, alternation from ear to ear at much faster or slower rates does not disrupt speech perception. Huggins (1964) elucidated these findings by showing that the rate of alternated speech that led to the poorest speech recognition was dependent on the speed of the speech signal. If speech is speeded up, faster alternation rates are necessary for minimal intelligibility.

By covarying the rate of alternation and the speed of the speech signal, Huggins found that alternating segments of speech approximating the average duration of a syllable led to the poorest recognition. Alternating speech at this rate insures that a switch between ears will occur exactly once during most syllables. Assuming that a speech signal in one ear cannot be integrated with the signal arriving in the opposite ear, these results indicate the syllable as an important segment in speech perception. With only part of the syllable in each ear, subjects would have minimal information if identification of the syllable was critical for speech perception. If the alternation of the speech was slowed down, the subject would have more of the syllable in one ear and, therefore, should be able to identify it more accurately. Since processing time is constant across rates of alternation in Huggin's study, the results indicate the syllable as the optimal segment or perceptual unit in the processing of speech.

Huggins (1964) also presents some measurements of the temporal characteristics of his speech passages. First, it is interesting that 18% of the total duration of the passages was essentially silent. This free time probably helps establish perceptual units and gives the subject time for perceptual processing. The mean duration of the syllables was 200 msec. with an interquartile range of 150–250 msec. The durations of syllables, then, agree with the estimates of perceptual processing time. The agreement of perceptual processing time and the duration of the perceptual unit correspond to a similar finding in visual information processing. Since each eye movement in reading overwrites the preceding visual image, the fixation must and does last long enough for perceptual processing (Haber, 1970; Woodworth, 1938).

The results indicate the syllable as the perceptual unit in speech perception. Warren (1970), in fact, has shown that if another extraneous sound replaces a phoneme in a sentence, listeners actually report hearing the missing phoneme. Warren deleted a 120-msec. section corresponding to the first /s/ in legislatures in the sentence, "The state governors met with their respective legislatures convening in the capital city." The missing section was replaced with a recorded cough or tone of the same duration. Only one of the 40 subjects reported hearing a missing sound and he selected the wrong one. In addition, subjects were not able to locate the position of the extraneous sound in the sentence. These results indicate that we can synthesize a unit of speech perception with a limited amount of information in the preperceptual auditory image.

The perceptual synthesis of the missing phoneme must be made on the cues from its auditory context. Relevant features exist before and after the missing phoneme and these features are held in preperceptual storage for perceptual processing. On the basis of limited information (since a phoneme is missing), the subject extracts what features are present and is able to synthesize the correct syllable or word. This synthesis must take place on a perceptual level rather than a verbal or abstract level since the subject actually believes he heard the missing phoneme. If the phoneme is replaced with silence rather than an extraneous noise, the subject notices the gap and yet is able to identify the word correctly. The temporal course of the information in the preperceptual store could be determined by varying the temporal separation between the degraded word and the necessary context for word synthesis. Phonemic restoration should not occur when the sensory features near the missing phoneme are no longer available for perceptual synthesis. Accordingly, if the context comes after this preperceptual information has decayed or has been masked by other sensory input, the observer should notice the missing phoneme.

In a different approach to disrupting speech, Cherry and Wiley (1967) and Holloway (1970) have shown that passing only the strongly voiced speech sounds decreases speech perception. This sequence of sounds and silent periods does not contain the rhythmic nature of normal speech. By adding a low level of white noise in the silent gaps between the separate sounds, the authors showed that the speech once again becomes intelligible. These results show that the nature of normal speech is important for establishing temporal or perceptual units for speech perception. By artificially segregating the information, the organization of the larger units normally identified is broken. Noise, then, can bridge the gap and provide the subject with an organized unit for speech perception.

Liberman, Cooper, Shankweiler, and Studdert-Kennedy (1967) have presented a detailed and interesting account of speech perception that differs from the present approach. The basic assumption of Liberman et al. is that speech perception occurs at the level of the phoneme. Therefore, their emphasis is on the information processing necessary for phonemic perception from the acoustic signal. The analysis of the acoustic signal is in terms of the formant structure of speech patterns. Figure 3 presents the first two formants of the patterns /di/ and /du/. The darkened areas represent the concentration of acoustic energy at the indicated frequencies over time.

Using the example of /di/ and /du/, Liberman et al. show that the second formant structure is necessary for perception of the /d/ segment of the syllables /di/ and /du/. Therefore, they argue that /d/ perceived as the same phoneme is represented by two entirely different acoustic signals (cf. Figure 3). Liberman et al. (1967) state,

> It is, then, the second-formant transitions that are, in the patterns in Figure 1 [Figure 3 in the present article], the acoustic cues for the perception of the /d/ segment of the syllables /di/ and /du/. We would first note that /d/ is the same perceptually in the two cases, and then see how different are the acoustic cues. In the case of /di/ the transition rises from approximately 2200 cps to 2600 cps; in /du/ it falls from about 1200 to 700 cps. In other words, what is perceived as the same phoneme is cued, in different contexts, by features that are vastly different in acoustic terms. How different these acoustic features are in nonspeech perception can be determined by removing them from the patterns of Figure 1 [Figure 3 in the present article] and sounding them in isolation. When we do that, the transition isolated from the /di/ pattern sounds like a rapidly rising whistle or glissando on high pitches, the one from /du/ like a rapidly falling whistle on low pitches [p. 435].

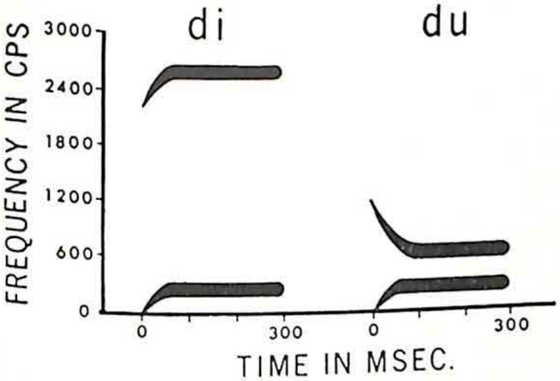


FIG. 3. Spectrographic patterns sufficient for the synthesis of /d/ before /i/ and /u/. (Reprinted with permission from an article by A. M. Liberman, F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy published in the *Psychological Review*, 1967, Vol. 74. Copyrighted by the American Psychological Association, Inc., 1967.)

However, assuming that the syllables /di/ and /du/ function as perceptual units, there exists a direct correspondence between acoustic signal and percept.

Other results of Liberman et al. (1967) indicate that presenting the early part of the pattern /di/ alone produces a nonspeech sound. By adding successive parts of the pattern, the phoneme /d/ is never heard. Rather the perception changes categorically to the syllable /di/. Therefore, with increases in stimulus duration, perception changes from a nonspeech sound to a consonant-vowel cluster. Liberman et al. conclude from this evidence that the formant transition continuously provides information about the two phonemes in parallel.

In contrast to the conclusion of Liberman et al., their results provide convincing support for eliminating the phoneme as the perceptual unit for processing speech. Evidence reviewed above indicated the syllable as the perceptual unit critical for speech perception. Other recent results show that phoneme perception follows the perception of the syllables containing the phonemes. Savin and Bever (1970) presented subjects with a sequence of nonsense syllables. The subjects responded as soon as they heard a target in the sequence. They responded faster to a target that was a complete syllable than to a target that was a phoneme from that syllable. Warren (1971) has studied identification time for targets in continuous speech. Identification times for monosyllabic words or nonsense syllables were faster than identification times for phoneme clusters. Reaction times to phoneme clusters were shorter than reaction times to individual phonemes within the clusters. These results indicate that syllable identification precedes phoneme identification in speech perception.

These results suggest a different process for syllable identification than the process responsible for phoneme identification. If the syllable is the perceptual unit, identification of the syllable involves an analysis of the information in the preperceptual storage of the syllabic unit. The distinctive features necessary for recognition would be relevant to this unit. On the other hand, since the phoneme is not a perceptual unit, phoneme identification in speech probably does not occur as a result of analysis of preperceptual memory. After identification of the syllable, the subjects could simply scan the syllable for the appropriate phoneme target. In vision, this process would be analogous to finding the letter "a" in the word "cat" after the word was already identified. Warren's (1970) results indicated that subjects do not recognize that a phoneme is missing when it is replaced by an extraneous sound. These results support the interpretation given here that the phoneme is not perceived directly from preperceptual memory, but is inferred from the identification of the syllable or word.

The results indicate perceptual units that are larger than the phoneme in speech perception. As Liberman et al. (1967) point out, attempts to produce intelligible speech by a recombination of phonemes has been unsuccessful (Harris, 1953). In contrast, Petersen, Wang, and Sivertsen (1958) showed that intelligible speech can be produced by using units that are at least one-half syllable in length. Liberman et al. (1967) concluded that

> This parallel delivery of information produces at the acoustic level the merging of influences we have already referred to and yields irreducible acoustic

> segments of approximately syllabic dimensions [p. 441].

Rather than parallel processing of phonemes, it is more parsimonious to assume that the syllable functions as the perceptual unit for speech perception.[5]

If the vowel is integrated with the consonant-vowel transition (as in Hirsh's task), the extended duration of the vowel in normal speech would provide time for perceptual processing. As mentioned earlier, vowels in normal speech last 150–250 msec. As can be seen in Figure 3, the last 200 msec. of /di/ and /du/ do not provide new information. The first 100 msec. of the pattern probably could be identified if it is presented alone and followed by a silent interval. However, if the consonant-vowel transition was followed by a speech sound that could not be integrated with it, perception should be disrupted and backward recognition masking should occur.

Throughout this paper it has been argued that there are perceptual or temporal units of auditory information, and that perception is determined by the processing of the information in the temporal units. The assumption of temporal units in auditory information processing may seem similar to the psychological moment theory proposed by Stroud (1955). However, the important difference is that the psychological moment assumes that time is the critical independent variable determining the perceptual experience (Kahneman, 1968; Schmidt & Kristofferson, 1963; Stroud, 1955). Contrasting this, the results indicate that the size of the temporal unit is determined by the properties of the physical stimulus. The duration of the perceptual unit in a consonant-vowel transition is longer than the 20-msec vowels used in the recognition paradigm. However, in both cases, we must process the information available and this processing requires time. During perceptual processing, the information is extracted from the stimulus image and the synthesized percept enters consciousness or short-term memory as a unit. The preperceptual features of the auditory input will usually be lost with the next auditory input. Therefore, short-term memory is necessary for retaining the synthesized information necessary for combination with the synthesized information from other perceptual units (Lashley, 1951; Wickelgren, 1969).

Auditory Short-Term Memory

Morton (1970) and Crowder and Morton (1969) interpret some short-term effects as evidence for a precategorical acoustic storage that is assumed to have many of the properties of an auditory preperceptual store. Morton and Crowder's demonstration of precategorical acoustic storage is the suffix effect. The subjects are asked to recall serially a list of eight items presented auditorily at a rate of two per second. Performance on a control list is compared to performance on the same list followed by the suffix zero. Although the suffix is not recalled, it selectively decreases performance on the final two or three items (Crowder, 1967). Control studies show that the auditory suffix zero produces more interference with recall of the final items than a visual suffix (Morton & Holloway, 1970).

These experiments indicate that although the suffix interferes with auditory information from the stimulus list, this information is *not* necessarily precategorical. It was argued earlier that the information available in the auditory image of a short tone was preperceptual since the information in the image was present before the tone had been identified. On the other hand, in the serial recall task, each item in the list presentation should have been identified before the following item is presented. Therefore, the suffix effect is due to interference of the subject's short-term memory for the final items in the list. Recognition masking, on the other hand, is due to interference with the subject's perception of the test item. If subjects were required to recall only the last two or three items in the list, the suffix effect should be eliminated. Since each item is perceived as it is presented, no interference should be observed in the recall task without some short-term memory loss.

---

[5] It should be noted that the term "syllable" must be interpreted loosely. There are four basic syllable types: Vs, VCs, CVs, and CVCs where V is a vowel and C is a single consonant or a consonant cluster. It is unlikely that all syllables function as perceptual units. For example, CVCs probably contain two perceptual units.

Crowder and Morton's results are more consistent with interference effects in short-term memory than precategorical acoustic storage. A number of results indicate that decreasing the similarity of the suffix to the stimulus list decreases interference. First, presenting the list in a male voice and the suffix in a female voice reduces the suffix effect. Second, presenting the suffix contralateral to the stimulus list presentation also reduces the suffix effect. Finally, white noise or a buzzer presented as the suffix produces no interference. Crowder and Morton's results, then, demonstrate that similarity plays an important role in short-term memory. This interpretation agrees with recent studies of short-term memory that find that increasing the similarity of retroactive items to test items increases forgetting (Deutsch, 1970; Massaro, 1970c, Experiment II; Reitman, 1971; Wickelgren, 1965, 1966).

The size of the suffix effect is directly related to the similarity of the stimulus list and the suffix. Contrasting this, similarity effects have not been found in studies of interference with preperceptual auditory images. Massaro (1970b) showed that decreasing the similarity between the test and masking tones did not decrease backward recognition masking. Furthermore, presenting the masking tone contralateral to the test tone presentation did not reduce the interference effect of the masking tone. Finally, white noise interferes with preperceptual memory in speech recognition (Dirks & Bower, 1970). These results support the assumption that the suffix effect and backward recognition masking occur at two different stages of information processing. Crowder and Morton are studying forgetting of identified items and have not demonstrated conclusively that memory for these items contains "precategorical acoustic storage."

As Morton (1970) points out, many of the results demonstrating a suffix effect can be interpreted in terms of selective attention. For example, presenting the suffix at twice the subjective loudness of the stimulus list reduces the suffix effect. In terms of selective attention, the difference in loudness of the suffix provides a dimension for rejection of the suffix. On the other hand, if precategorical acoustic storage was responsible for the suffix effect, a louder suffix would be expected to produce more interference, since the suffix supposedly overwrites the information in precategorical storage. Morton (1970) presents the crucial test between selective attention and precategorical acoustic storage accounts of the suffix effect. Presenting the stimulus list monaurally, he compared a monaural to a binaural presentation of the suffix. Selective attention would predict a reduction in the suffix effect with the binaural presentation, since this could be used as a dimension for rejection. However, "precategorical acoustic storage" predicts no difference, since a binaural suffix should overwrite preperceptual information as well as a monaural suffix. The results provide evidence against a "precategorical acoustic storage" account of the suffix effect, since the binaural suffix actually produced less interference than the monaural suffix.

Bryden (1971), working in a dichotic listening situation, has also interpreted attention and short-term memory effects as preperceptual auditory memory effects. Bryden presents four pairs of digits, at two pairs per second, in a dichotic listening experiment, The subject is instructed to attend to a given channel. He then recalls all of the digits with instructions to recall the attended or unattended message first. Results indicate that, overall, the attended digits are recalled much better than the unattended digits. Furthermore, an analysis of the serial position curve indicates that there is a very small serial position effect for the attended digits. In contrast, memory for the unattended digits shows a huge recency effect. Recall of the last unattended digit is twice as probable as recall for the first unattended digit.

Bryden (1971) interprets these results as indicating that the unattended items are in preperceptual storage. However, the digits presented in the unattended ear could have been identified at presentation. Supporting this interpretation, Norman (1969) has shown that items presented to the unattended ear in a shadowing task are recognized at presentation. These unattended items also show forgetting functions similar to typical short-term memory studies (Norman, 1966). Therefore, it is argued that the unattended items in Bryden's task were identified at presentation but received very little perceptual processing for memory and, therefore, showed more forgetting than the attended digits.

This interpretation requires a distinction between two stages of perceptual processing pointed out explicitly by Massaro (1970a). The first stage, perception or perceptual processing of the information in the preperceptual auditory image refers to the analysis of the sensory input in which physical features are examined in order to identify the input. After identification of the item, the second stage of perceptual processing is carried out in order to remember the item. In Bryden's task the unattended items could have been poorly recalled because they received very little perceptual processing for memory. Digits can be identified easily within 200 msec. as shown by Warren and Warren (1970). Therefore, Bryden's subjects had time to identify both the attended and unattended items at presentation. In the time remaining, they rehearsed the attended items. This interpretation is more compatible with the literature reviewed here than is an interpretation based on preperceptual storage.

Preperceptual auditory images and short-term auditory memory have their analogous counterparts in vision. A preperceptual auditory image is analogous to the positive afterimage in vision that keeps the stimulus in its original form (Sperling, 1960). Evidence presented earlier indicated that very little preperceptual information remains after readout of the auditory image which occurs within 250 msec. Accordingly, synthesized short-term auditory memory must be qualitatively different than preperceptual auditory images. Synthesized auditory memory is analogous to the visual memory studied by Posner and his colleagues (Beller, 1971; Posner, Boies, Eichelman, & Taylor, 1969; Posner & Keele, 1967). Their results indicate that the visual memory used to remember a capital A is qualitatively different than preperceptual visual images. Similarly, the auditory memory used to remember a list of digits is qualitatively different than preperceptual auditory images.

Conclusion

Temporal or perceptual units play an important role in auditory recognition. These units are held in a preperceptual auditory store for perceptual processing. If a second auditory pattern can be integrated with a first, they can form a single unit. If not, the second can interfere with the preperceptual image of the first. Perceptual processing refers to an analysis of information in the perceptual unit. This analysis requires an examination of the physical features of the stored sequential pattern in order to identify the input. The temporal course of perceptual processing depends on the complexity of the identification task. The more difficult the discrimination, the longer the time needed for reading out the necessary information.

The present results are also relevant to the distinction drawn between the processing of speech and nonspeech material (Liberman et al., 1967). The motor theory of speech perception assumes that articulatory mechanisms mediate the perceptual processing of speech. However, assuming a preperceptual auditory store eliminates some of the "complexities" of speech input that demand a unique processor for speech. Most importantly, the perceptual unit of speech cannot be at the phoneme level but must be more complicated. On the other hand, it cannot be as large as two or three words as noted in the discussion of short-term memory experiments. Phonemes may become important when the features necessary for recognition are determined. However, this is an entirely different problem than determining the perceptual unit that remains intact for perceptual processing.

The present paper has not attempted to provide a total account of speech perception or complex pattern recognition where short-term memory, long-term memory, and decision processes become critical. With respect to speech perception, the contributions of syntax and semantics cannot be ignored (Miller, 1962; Miller, Heise, & Lichten, 1951). With complex auditory patterns, form, rhythm, melody, and organization become important at the level of synthetic memory (Cheatham & White, 1954; Dowling, 1971; Garner, 1951; Heise & Miller, 1951; White, 1963). Enjoying a symphony involves much more than simply analyzing the information available in preperceptual store. The present paper has, however, provided a conceptual framework for the recognition stage of perceptual processing. This stage of perceptual processing outputs a syntheiszed percept that becomes available to the next stage of auditory information processing.

REFERENCES

Averbach, E., & Coriell, A. S. Short-term memory in vision. *Bell Systems Technical Journal*, 1961, **40**, 309–328.

Beller, H. K. Priming: Effects of advance information on matching. *Journal of Experimental Psychology*, 1971, **87**, 176–182.

Broadbent, D. E., & Ladefoged, P. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, **31**, 1539.

Byrden, M. P. Attentional strategies and short-term memory in dichotic listening. *Cognitive Psychology*, 1971, **2**, 99–116.

Cheatham, P. G., & White, C. T. Temporal numerosity: III. Auditory perception of number. *Journal of Experimental Psychology*, 1954, **47**, 425–432.

Cherry, E. C. Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 1953, **25**, 975–979.

Cherry, E. C., & Taylor, W. K. Some further experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 1954, **26**, 554–559.

Cherry, C., & Wiley, R. Speech communication in very noisy environments. *Nature*, 1967, **214**, 1164.

Creel, W., Boomsliter, P. C., & Powers, S. R. Sensations of tones as perceptual forms. *Psychological Review*, 1970, **77**, 534–545.

Crowder, R. G. Prefix effects in immediate memory. *Canadian Journal of Psychology*, 1967, **21**, 450–461.

Crowder, R. G., & Morton, J. Precategorical acoustic storage. *Perception and Psychophysics*, 1969, **5**, 365–373.

Deatherage, B. H., & Evans, T. R. Binaural masking: Backward, forward, and simultaneous effects. *Journal of the Acoustical Society of America*, 1969, **46**, 362–371.

Deutsch, D. Tones and numbers: Specificity of interference in immediate memory. *Science*, 1970, **168**, 1604–1605.

Dirks, D. D., & Bower, D. Effect of forward and backward masking on speech intelligibility. *Journal of the Acoustical Society of America*, 1970, **47**, 1003–1008.

Donders, F. C. On the speed of mental processes. (Trans. by W. G. Koster) In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North Holland Publishing, 1969.

Dowling, W. J. Recognition of inversions of melodies and melodic contours. *Perception and Psychophysics*, 1971, **9**, 348–349.

Efron, R. Effect of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics*, 1970, **8**, 231–234. (a)

Efron, R. The minimum duration of a perception. *Neuropsychologia*, 1970, **8**, 57–63. (b)

Efron, R. The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia*, 1970, **8**, 37–55. (c)

Elliot, L. L. Backward and forward masking of probe tones of different frequencies. *Journal of the Acoustical Society of America*, 1962, **34**, 1116–1117. (a)

Elliot, L. L. Backward masking: Monotic and dichotic conditions. *Journal of the Acoustical Society of America*, 1962, **34**, 1108–1115. (b)

Elliot, L. L. Development of auditory narrow-band frequency contours. *Journal of the Acoustical Society of America*, 1967, **42**, 143–153.

Erikson, C. W., & Johnson, H. J. Storage and decay characteristics of nonattended stimuli. *Journal of Experimental Psychology*, 1964, **68**, 28–36.

Fletcher, H. *Speech and hearing in communication*. Princeton: Van Nostrand, 1953.

Galanter, E. Contemporary psychophysics. In, *New directions in psychology*. New York: Holt, Rinehart & Winston, 1962.

Garner, W. R. The accuracy of counting repeated short tones. *Journal of Experimental Psychology*, 1951, **41**, 310–316.

Gol'dburt, S. N. Investigation of the stability of auditory processes in micro-intervals of time (new findings on back masking). *Biofizika*, 1961, **6**, 717–724. (English translation: *Biophysics*, 1961, **6**, 809–817.)

Gray, G. W. Phonemic microtomy: The minimum duration of perceptible speech sounds. *Speech Monographs*, 1942, **9**, 75–90.

Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966.

Guttman, N., & Julesz, B. Lower limits of auditory periodicity analysis. *Journal of the Acoustical Society of America*, 1963, **35**, 610.

Haber, R. N. (Ed.) *Information-processing approaches to visual perception.* New York: Holt, Rinehart & Winston, 1969.

Haber, R. N. How we remember what we see. *Scientific American*, 1970, **222** (2), 104–112.

Haber, R. N., & Standing, L. G. Direct measures of short-term visual storage. *Quarterly Journal of Experimental Psychology*, 1969, **21**, 43–54.

Harris, C. M. Study of the building blocks in speech. *Journal of the Acoustical Society of America*, 1953, **25**, 962–969.

Heise, G. A., & Miller, G. A. An experimental study of auditory patterns. *American Journal of Psychology*, 1951, **64**, 68–77.

Hirsh, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, **31**, 759–767.

Hochberg, J. Attention, organization and consciousness. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis.* New York: Appleton-Century-Crofts, 1970.

Holloway, C. M. Passing the strongly voiced components of noisy speech. *Nature*, 1970, **226**, 178–179.

Homick, J. L., Elfner, L. F., & Boothe, G. G. Auditory temporal masking and the perception of order. *Journal of the Acoustical Society of America*, 1969, **45**, 712–718.

House, A. S. On vowel duration in English. *Journal of the Acoustical Society of America*, 1961, **33**, 1174–1178.

Huggins, A. W. F. Distortion of the temporal pattern of speech: Interruption and alternation. *Journal of the Acoustical Society of America*, 1964, **36**, 1055–1064.

Jakobson, R., & Halle, M. *Fundamentals of language.* 's-Gravenhage: Mouton, 1956.

Kahneman, D. Method, findings, and theory in the studies of visual masking. *Psychological Bulletin*, 1968, **70**, 404–425.

Koffka, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley, 1951.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, **74**, 431–461.

Massaro, D. W. Perceptual processes and forgetting in memory tasks. *Psychological Review*, 1970, **77**, 557–567. (a)

Massaro, D. W. Preperceptual auditory images. *Journal of Experimental Psychology*, 1970, **85**, 411–417. (b)

Massaro, D. W. Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1970, **83**, 32–39. (c)

Massaro, D. W. Effect of masking tone duration on preperceptual auditory images. *Journal of Experimental Psychology*, 1971, **87**, 146–148.

Massaro, D. W. Stimulus information versus processing time in auditory pattern recognition. *Perception and Psychophysics*, 1972, in press.

McGill, W. J. Loudness and reaction time. *Acta Psychologica*, 1961, **19**, 193–199.

McGill, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963.

Miller, G. A. Decision units in the perception of speech. *IRE Transactions in Information Theory*, 1962, IT-8, 81–83.

Miller, G. A., Heise, G. A., & Lichten, W. The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 1951, **41**, 329–335.

Miller, G. A., & Licklider, J. R. Intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 1950, **22**, 167–173.

Miller, G. A., & Taylor, J. Perception of repeated bursts of noise. *Journal of the Acoustical Society of America*, 1948, **20**, 171–182.

Morton, J. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

Morton, J., & Holloway, C. M. Absence of a cross-modal "suffix effect" in short-term memory. *Quarterly Journal of Experimental Psychology*, 1970, **22**, 167–176.

Neisser, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

Norman, D. A. Acquisition and retention in short-term memory. *Journal of Experimental Psychology*, 1966, **72**, 369–381.

Norman, D. A. Memory while shadowing. *Quarterly Journal of Experimental Psychology*, 1969, **21**, 85–93.

Patterson, J. H., & Green, D. M. Discrimination of transient signals having identical energy spectra. *Journal of the Acoustical Society of America*, 1970, **48**, 894–905.

Peterson, G. E., Wang, W. S.-Y., & Sivertsen, E. Segmentation techniques in speech synthesis. *Journal of the Acoustical Society of America*, 1958, **30**, 739–742.

Plomp, R. Decay of auditory sensation. *Journal of the Acoustical Society of America*, 1964, **36**, 277–282.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. I. Retention of visual and name codes of single letters. *Journal of Experimental Psychology*, 1969, **79** (1, Pt. 2).

Posner, M. I., & Keele, S. W. Decay of visual information from a single letter. *Science*, 1967, **158**, 137–139.

Raab, D. H. Backward masking. *Psychological Bulletin*, 1963, **60**, 118–129.

Reitman, J. S. Mechanisms of forgetting in short-term memory. *Cognitive Psychology*, 1971, **2**, 185–195.

Samoilova, I. K. Masking of short tone signals as a function of the time interval between masked and masking sounds. *Biophysics*, 1959, **4**(5), 44–52.

Savin, H. B., & Bever, G. T. The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 295–302,

Schmidt, M. W., & Kristofferson, A. B. Discrimination of successiveness: A test of a model of attention. *Science*, 1963, **139**, 112–113.

Sperling, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, **74**(11, Whole No. 498).

Sperling, G. Successive approximations to a model for short-term memory. *Acta Psychologica*, 1967, **27**, 285–292.

Sternberg, S. The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, 1969, **30**, 276–315.

Stevens, S. S., & Davis, H. *Hearing*. New York: Wiley, 1938.

Stroud, J. The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology*. New York: Free Press, 1955.

Studdert-Kennedy, M., Shankweiler, D., & Schulman, S. Opposed effects of a delayed channel on perception of dichotically and monotically presented CV syllables. *Journal of the Acoustical Society of America*, 1970, **48**, 599–602.

Thomas, I. B., Hill, P. B., Carroll, F. S., & Bienvenido, G. Temporal order in the perception of vowels. *Journal of the Acoustical Society of America*, 1970, **48**, 1010–1013.

Von Békésy, G. Auditory inhibiton in concert halls. *Science*, 1971, **171**, 529–536.

Warren, R. M. Perceptual restoration of missing speech sounds. *Science*, 1970, **167**, 392–393.

Warren, R. M. Identification times for phonemic components of graded complexity and for spelling of speech. *Perception and Psychophysics*, 1971, **9**, 345–349.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusion of patterns other than speech or music. *Science*, 1969, **164**, 586–587.

Warren, R. M., & Warren, R. P. Auditory illusions and confusions. *Scientific American*, 1970, **223**(6) 30–36.

White, C. T. Temporal numerosity and the psychological unit of duration. *Psychological Monographs*, 1963, **77**(12, Whole No. 575).

Wickelgren, W. A. Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 53–61.

Wickelgren, W. A. Phonemic similarity and interference in short-term memory for single letters. *Journal of Experimental Psychology*, 1966, **71**, 396–404.

Wickelgren, W. A. Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, 1969, **76**, 1–15.

Woodworth, R. S. *Experimental psychology*. New York: Holt, 1938.

(Received July 12, 1971)

Psychological Review
1972, Vol. 79, No. 2, 146–160

A MODIFIED BELIEF THEORY OF PREJUDICE EMPHASIZING THE MUTUAL CAUSALITY OF RACIAL PREJUDICE AND ANTICIPATED BELIEF DIFFERENCES[1]

RICHARD A. DIENSTBIER[2]

University of Nebraska-Lincoln

The belief theory of prejudice introduced by M. Rokeach stated that racial prejudice is the result of the anticipation of belief differences. The unidirectional causal relationship implied is criticized as oversimplified. Research supporting the belief theory is examined, with conceptual and experimental deficiencies noted. A new formulation is proposed which emphasizes mutual causality between racial prejudice and anticipated belief differences. Two studies supporting that view are presented in which belief communications were presented as tape-recorded interviews or speeches, with the race and social class of the communicator first having been manipulated. The interrelationships between communicator's race, specific communication topic, and subject's prejudice level on the dimensions of felt similarity of the subject to the communicator are seen as supporting the mutual causation formulation.

The belief theory of prejudice suggests that prejudice results from an aversion toward individuals who hold beliefs which are incongruent with one's own. It was postulated that the aversion toward incompatible beliefs was due to universal needs for consensual validation (Rokeach, Smith, & Evans, 1960) and/or the desire for the reinforcement provided by belief agreement (Byrne & Wong, 1962). Furthermore, belief theory proponents have suggested that without anticipations about belief differences, no prejudice would exist. Rokeach et al. (1960), in elaborating that point in the original formal presentation of the belief theory, asked and answered the ultimate question:

> Are there two qualitatively different kinds of prejudice, or can racial and ethnic prejudice be subsumed under belief prejudice? The data, such as they are, seem to yield up a primarily no answer to the first part, and primarily yes answer to the second part of the question [p. 157].

Since that statement, a good many other researchers have interpreted their research as supporting those notions.

The purposes of this paper are several: First, to demonstrate that the question of race versus belief cannot be answered in the abstract, and that any attempt at a quantitative answer to the question of the relative strength of race and belief differences in determining prejudice is inappropriate. In order to develop the argument leading to that conclusion, a review of the literature supporting the belief theory will be presented, discussed, and criticized. The second purpose of this paper is to show that the interaction of belief differences and race are quite complex and that a reformulation of the belief theory in terms of mutual causal relations between racial prejudice and beliefs (rather than the unidirectional relation presently in vogue) is more appropriate. Evidence in the form of data from two studies will be presented to demonstrate that race prejudice can exist regardless of belief differences, and that a more complex formulation than the belief theory is indicated.

---

[1] Study I of this paper was based on dissertation research submitted to the Department of Psychology, University of Rochester, in partial fulfillment of the requirements for the PhD degree. Special thanks are due to my thesis advisor, Vincent Nowlis, for assistance with the research and for comments offered concerning the manuscript, and to my wife, the interviewer in the research. Thanks are also due Allen Adinolfi, James Barron, Ed Fougerousse, and John Martinelli for help with the audio and video work.

Study II was made possible through a Junior Faculty Summer Research Fellowship provided by the Research Council of the University of Nebraska-Lincoln and through equipment funds provided by the Research Council. Thanks are due to Pamela Munter and Rick Scheidt for assistance in developing the script and tapes used in that study, to Garret Fox for assistance with the statistical analyses, and to Monte Page for his critical reading of the manuscript.

[2] Requests for reprints should be sent to Richard A. Dienstbier, Department of Psychology, University of Nebraska, Lincoln, Nebraska 68508.

BELIEF THEORY RESEARCH

In the initial belief theory research of Rokeach et al. (1960), stimulus individuals were presented in a factorial design of Race × Belief, so that the stimulus persons were presented as (*a*) Negro person who agrees; (*b*) Negro person who disagrees; (*c*) white person who agrees; (*d*) white person who disagrees. Agreement was manipulated by the stimulus person's stand on some one topic, so that most stimulus person presentations were as short as "A Negro who believes in God." Topics were both general and race relevant, including socialized medicine, communism, labor unions, God, immediate desegregation, fraternity and sorority integration, equality of race, and integration of neighborhoods. Prejudice was determined on the basis of the subject's responses to a 9-point social-distance-type scale concerning potential friendship. Using a factorial design of Race × Agreement (Belief), the data indicated that the basic hypothesis concerning the greater power of the belief manipulation was correct.

Almost identical stimulus person presentation methods were used by Smith, Williams, and Willis (1967) with subject populations different in location and age from those of the Rokeach research. Their findings generally tended to replicate the Rokeach findings, although a high degree of racial (as opposed to belief) discrimination was found for the Louisiana subject sample. There were, however, several significant problems with the research. The questionnaire was undisguised so that it was completely obvious to subjects that race and belief issues were being studied and that their responses would indicate both their attitudes toward the belief issues and toward race. Results must be interpreted, therefore, with the recognition of the confounding of social desirability with racial prejudice mediated by the obvious demand characteristics of the situation. Second, in this rather artificial situation, subjects must make some decisions as to what the experimenter means to imply by his stimulus materials. Does the mention of one man as Negro and another as white imply that all other things are equal between the two? Although a subject could blithly develop and accept such a premise in the context of psychological research, he might never be willing to concede the same point in real-life racial interactions. This objection also concerns the question of whether the laboratory presentation of race-relevant materials results in subjects believing the experimenter. Suspiciousness checks have been conspicuously absent from the bulk of the cited research. One does not read any reports of such direct questions as "Did you think we were looking at the influence of race on your responses?" A rare reported occurrence of a check of any sort (Hendrick, Bixenstine, & Hawkins, 1971) asked subjects "to describe their reactions to the experiment." Work on postexperimental questionnaire procedures for deception experiments has suggested the advantages of direct and specific questions following more general ones as an ideal technique for detecting true suspicions without eliciting false-positive reports (Page, 1971). Whether we consider the possibility of mistaken assumptions on the part of the subjects, or their possible suspiciousness of the experimenter's true intent, there is cause for serious doubt that subjects respond to the stimulus materials in a manner which is similar to their real-life behavior.

A second category of objection to the Rokeach et al. (1960) and Smith et al. (1967) studies was explicated by the research of Triandis and Davis (1965). Using stimulus person presentation techniques similar to the Rokeach et al. (1960) method, those authors presented to each of 300 subjects eight stimulus persons representing all combinations of race (white or Negro), sex, and strong pro or anti views toward civil rights. Stimulus persons were described in short phrases such as "Male, Negro, favors strong civil rights legislation." Dependent measures of semantic differential and social-distance-like behavioral differential scales were used. From previously determined factor structures (Triandis, 1964), the behavioral differential scales were classified into the five factors of formal social acceptance, friendship acceptance, marital acceptance, social distance, and subordination. For the first two factors, belief differences provided the greatest influence, while sex and race had the most impact on the third, and the last two were most influenced by race. These results were interpreted as suggesting that the dependent measures used determine the relative influence of the race versus belief influence. Other research (Dienstbier, 1970) extended those findings. For the same subjects, while favoritism may be indicated toward minority group members (over whites) on one scale such as "would elect to political office," significant negative prejudice may still be indicated on other more race-sensitive scales such as "accept as kin by marriage." Mezei (1971) has since suggested that the greater influence of race on such items as marriage acceptance is merely due to subjects' anticipation of social disapproval for a more lenient stand. The Mezei data are, unfortunately, not free from the alternative interpretation that race prejudice is very real, and that subjects see other people as sharing their prejudice in order to defend their own views and because (as the belief theory suggests) they want to believe that others largely share their beliefs and values. At present, then, it appears that certain dependent measures are far more responsive to racial prejudice while others are more belief sensitive. The friendship acceptance scales of the Rokeach et al. (1960) and the Smith et al. (1967) research and the choice of work partner scales used in Byrne's research (reviewed below) are far more susceptible to belief than race influences (Dienstbier, 1970; Triandis, 1961; Triandis & Davis, 1965). Their use, therefore, biases results toward confirmation of the belief theory.

Using a different stimulus person presentation technique, Byrne and Wong (1962) presented stimulus persons through a 26-item checklist of attitudes, with the valence of the attitudes of the stimulus persons having been derived from previously acquired self-reports of the subjects on those same attitudes. The stimulus persons' attitude checklists were constructed so that attitudes were like those of the individual subject (all responses to each item on the same side of the 6-point scale, but not necessarily in the same exact scale position as chosen by the subject), or unlike those of the subject. The checklist of attitudes developed by Byrne concerned such topics as marriage, entertainment, religion, politics, drinking, and a Catholic president; the items do *not*, according to Byrne and Wong, "reflect common elements in the Negro stereotype." Analysis of the factorial design of two levels of subject prejudice by two levels of race (Negro and white) by two levels of agreement (belief) found agreement to be the greatest source of variance on the dependent measures—choice of work partner and personal feeling scales.

In subsequent related research by Byrne and Mc Graw (1964), Stein, Hardyck, and Smith (1965), Stein (1966), Insko and Robinson (1967), and Robinson and Insko (1969), the attitude checklist belief manipulation method was elaborated and applied to different subject populations, in some cases with the addition of such factors as age and religious differences attributed to the stimulus individuals. The Byrne and Mc Graw (1964) research used the same checklist of attitude manipulations of belief as the previous Byrne research. Stein et al. (1965) and Stein (1966) used a teenager questionnaire of about 25 items, a few of which were relevant to stereotypes of blacks (concerning intelligence, morality, ambition, and dancing ability). Insko and Robinson (1967) and Robinson and Insko (1969) used many of the same belief manipulation items of Stein's research, adding several items taken from the Blake and Dennis (1943) study on Negro stereotypes. Robinson and Insko (1969) provided a clever refinement on the practice of making the dissimilar checklist by filling out the stimulus person's checklist with items scored opposite from the subject's. Their method involved finding what the subject's estimates of the typical black's attitudes, and returning that profile back to the subject later, attributed once to a white and once to a black individual. Although these studies generally attempted to disguise the true nature of the research by suggesting to the subjects that the stimulus individuals were real people, the attitude checklist technique still suffers under the obvious limitation of being quite artificial, so that applicability of results to behavior in nonlaboratory situations is quite limited.

These studies conceptualized and operationalized prejudice in a variety of ways. Although Byrne and Mc Graw (1964) used only the choice of friendship and work partner measures of the previous Byrne research, the research of Stein (1966) and Stein et al. (1965) used social-distance-type measures which ranged into such race-sensitive areas as acceptance as kin by marriage. The dependent measures of Insko and Robinson (1967) and Robinson and Insko (1969) included both semantic differential and social-distance-type items.

Generally the Byrne and Mc Graw (1964) research supported the belief theory predictions, as did the Stein et al. (1965) and Stein (1966) research (on all prejudice measures except that of "accept as kin by marriage" used in the Stein et al., 1965, research). The southern samples of the Insko and Robinson (1967) and the Robinson and Inkso (1969) studies, however, indicated a greater degree of race than belief prejudice on most of the social-distance-type measures which were based on Triandis' behavioral differential factors of friendship acceptance and social distance. Semantic differential scales generally indicated greater belief than race effects (Insko & Robinson, 1967; Robinson & Insko, 1969).

Other experiments, which were designed to test the belief theory but which used stimulus person presentation techniques different from those described above, did not provide uniform support for the belief theory. Triandis (1961) presented stimulus individuals by religious, occupational, and racial labels, asking each subject to imagine the stimulus individual to be of the same or different philosophy as the subject. Results did not support the belief theory, instead they indicated race to be the most important variable in determining prejudice on a variety of social distance scales (in contrast to similarity of philosophy, religion, and occupation). In three separate experiments, Rokeach and Mezei (1966) had subjects interact with four confederates, two of whom (one white and one Negro) agreed with the subject and two of whom disagreed (on topics not directly relevant to race or race-stereotype issues). Using the subject's choice of two coffee partners or two work partners as a measure of attraction toward the confederates, the experimenters were able to determine that more choices were made on the basis of belief alone than of race alone. Frequent choices of partners in combinations which could not be attributed to race or belief similarity, however, limit the degree to which those results can be interpreted as supportive of the belief theory. In a study similar in conception to that of Rokeach and Mezei (1966), Hendrick et al. (1971) presented subjects with a videotape sequence of two whites and two blacks discussing the issue of the Vietnam war. Using measures of felt similarity, liking, and trait ratings as prejudice indexes, the authors found that the actors' race had much weaker effects than their position on the war issue, concluding: "The results of the study provide substantial support for the Rokeach et al. theory [p. 255]." "*In fact, race may not be a very important variable at all* [italics added, p. 257]" in determining attraction. The authors note, however, that those conclusions do not extend to such measures as the "date my sister" social distance item used in their research (which showed large race effects).

Although the Rokeach and Mezei (1966) and Hendrick et al. (1971) studies used more lifelike presentations of the stimulus individuals and their communications than did the other reviewed studies, Rokeach and Mezei (1966) used dependent measures which are more sensitive to belief manipulations, as described above, and both studies used communications devoid of race-relevant content. With respect to this last issue, there is little value in manipulating belief issues which have nothing to do with racial stereotypes and prejudice. If white Americans feel antipathy toward black Americans based on assumptions about the blacks' beliefs, those beliefs will not as likely be beliefs about the Vietnam war as about ambition, morality, cleanliness, dependability, etc. The race prejudice reduction achieved in the several studies which have used belief issues irrelevant to racial prejudice and stereotypes have probably achieved that reduction partly because student subjects are generally somewhat reluctant to make responses which appear prejudiced, partly because their initial level of race prejudice might be very low, and partly because they may make unwarranted assumptions about the stimuli toward which they should respond.

The differences in findings between researchers using different techniques to manipulate independent variables naturally provide the stuff of controversy. Logically, there appear to be an unlimited number of ways in which either race or belief information could be manipulated, making dubious the conclusions of studies which purport to show that one class of variables is more powerful than the other. Some researchers in the area have been aware of this problem, though such awareness has not always been apparent in titles which have ranged from "Two kinds of prejudice or one?" (Rokeach et al., 1960), and "Race and belief: An open and shut case" (Stein et al., 1965), to "Race versus belief similarity as determinants of attraction: A search for a fair test" (Hendrick et al., 1971). Responding to the logical demands of the situation, Byrne (see Byrne & Irwin, 1969), who had himself once been in the thick of the controversy, concluded that the question of relative strength could not be meaningfully answered. As indicated in the review above, however, the debate was subsequently taken up by Hendrick et al. (1971), who concluded that finally they had the "fair test," so that "while assessment of relative power is difficult, the question is certainly not meaningless." This test was achieved through "approximately equating . . . intuitively" the relative strength of the two independent variables.

ALTERNATE FORMULATIONS TO THE BELIEF THEORY

Although the logical problem of the relative strength of factorially presented race and belief manipulations is serious, other aspects of the belief theory and its supporting research are also weak. The explication of these problems will be aided by the consideration of a more complex view of prejudice presented by Allport (1954) six years before the advent of the belief theory. Allport defined prejudice as "an antipathy based upon a faulty and inflexible generalization." Within the definition itself, there is no clue as to the origins of that antipathy, but Allport suggested that the bases are legion, including fear of strangeness, felt needs to justify discriminatory or economic practices, various ego-defense needs often associated with guilt, and ego-defense functions of projection and displacement. Often too, prejudice was conceived by Allport to result from the child simply imitating the antipathies of the parents. It is apparent that Allport's understanding of prejudice is not completely harmonious with belief theory notions. Whereas Allport suggested that the *negative beliefs about an ethnic group might result from negative feelings* which themselves may exist for any number of reasons, Rokeach et al. (1960) suggest rather unequivocally that the causal sequence is the opposite—*prejudice results from perceived belief differences.* Thus, while the first view sees prejudice as having complex causal roots, with feelings and beliefs being complexly related in mutually supportive relationships, the belief view attempts to explain prejudice on a far simpler basis and implies a simpler view of man.

The unidirectional belief theory view of "prejudice from anticipated belief differences" implies that information presented to correct belief misconceptions should result in prejudice being monotonically reduced. But social observation suggests that this may not occur. There are numerous examples of minority populations existing for years within majority cultures with no changes in the majority's often inaccurate stereotypes. One example of this phenomenon is the persistence of the myth of Negro male sexual prowess so popular in the South (Dollard, 1937). It seems, then, that some erroneous beliefs that are the cause of racial prejudice (to follow the causal relationships espoused by the belief theories) are not easily changed by contradictory day-to-day evidence. This suggests that these beliefs, being especially resistant to change, might serve some psychological needs of their own—that they might result from needs to bolster and justify prejudice (affect) and discrimination which, in turn, might serve a multitude of economic, political, personality, and/or social status needs (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; Allport, 1954).

The question of which comes first, the affect (prejudice) or the cognitive support (stereotype, or beliefs about beliefs) is a debatable issue, but might best be resolved by admitting to mutual causal relationships between the variables, especially, of course, if those mutual relationships can be demonstrated. It is curious that this type of controversy develops in modern psychology. In the abstract, most of us agree that most psychological processes conceptualized at the level of cognition and affect seldom occur in unidirectional causal relationships, but involve instead complex feedback and interaction. Often, however, when a theorist demonstrates one side of that loop as his area of special interest, he is engaged in debate by others who choose to emphasize the opposite part of the loop; this either-or-type debate tends to becloud the issues. To return to the specific issue, it may, indeed, be quite appropriate to suggest that the anticipation of belief differences causes prejudice, but if we were to stop there, other important and complicating relationships might well be missed. Data presented in this paper indicate that prejudice causes the anticipation of belief differences. That opposite causal relationship does not negate the belief view; we should visualize a relationship of mutual influence between antipathy and expectancies about beliefs.

A THEORETICAL RESOLUTION

The theoretical position of this paper, then, is as follows: While research supporting the belief theory indicates that prejudice is based on or caused by the anticipation of belief differences, certain real-life observations and other psychological evidence suggest that some other complicating factors influence the dynamics of prejudice. The research reviewed above, which has often been cited as providing apparent evidence for the belief theory (Insko, 1967; Kirscht & Dillehay, 1967), fails to withstand this argument both on the logical grounds explicated above and because certain critical complicating factors, present in real life, have been absent from those research efforts.

It could be argued that the intent of all psychological laboratory research is to isolate certain critical variables from the complications of real life, and that our conclusions are always tempered by an implied "all other things being equal." This is a justifiable approach, but it requires that when we interpret the significance of our findings for the real world, we temper them with the qualifications that they have come from relatively artificial and contrived situations. This tempering has not been applied to the belief theory research. The belief theory researchers, having investigated prejudice from few limited paradigms, have stacked the experimental deck to prove that belief differences account for it all. The two experiments presented in this paper are an attempt to demonstrate how one might "unstack" such experiments by using belief information which is very relevant to race, by presenting that information in a more real-life-like context, and by using dependent measures which provide a broad range of prejudice indexes. Even when all this is done, of course, it is still illogical to try to assess the relative influence of race and belief.

The hypothesis underlying the research of this paper is simply that it is not possible to consider *realistic and racially relevant* belief manipulations to be independent of the race manipulations; these two types of variables are, when realistically presented, interactive. And that interaction is extremely (perhaps hopelessly) complex. As information is received, the degree it will be differentially interpreted by different listeners will depend on the manner of information presentation (is interpretation possible, or is a checklist of attitudes provided), the specific nature of the information (does the listener have race-relevant preconceptions concerning this specific topic), the salience of the race differences (did the term "Negro" appear on a form, or has one seen the individual or heard his voice), and the prejudice level of the listener. The research discussed in this paper was designed to illustrate part of this complexity—to demonstrate that a communicator's race and communicated information do interact, or that the manipulation of race influences the interpretation of information.

Consider again the research which has been applied to the question of race versus belief. The two paradigms which have tended to consistently support the belief theory (the simple statements of Rokeach et al., 1960, and the attitude checklist technique originated by Byrne & Wong, 1962) both used belief presentation techniques which permitted the subjects very limited latitude for interpretation. Had that information been presented in a more lifelike or equivocal form, it would have been more possible for the subjects' race prejudices to have influenced their understanding or interpretation of the information, with the result that relatively greater proportions of the variance in those studies would have been contributed by race variables. The possibility that the results might be quite different when information presentation techniques are different, perhaps more lifelike, will become more evident with the presentation of the data of this paper.

RESEARCH SUPPORTING A MORE COMPLEX RELATIONSHIP

These data were derived from two separate research efforts, the first conducted with middle-class women in Rochester, New York, the second with university students in Lincoln, Nebraska. Although the first study was a complex study concerning a variety of questions pertaining to prejudice, only that portion of its data relevant to the present issue and to the Lincoln study is discussed.

The hypothesis of the first study was that if an individual were overheard talking about himself, and about his beliefs, the degree to which subjects would rate their beliefs as similar to the speaker's would depend on both the apparent race and class of the speaker. Since the results of that research influenced the development of the second study, a preview of the results of the first is necessary here. While lower-class status of the communicator resulted in the subjects' seeing their views as being less similar to the communicator across most belief issues discussed, race had a much more variable effect, obviously interacting with the specific belief issue, sometimes influencing the listener toward greater agreement with the black speaker (compared to the white communicator), but sometimes influencing in the opposite direction. This finding, that agreement with the stimulus individual appears to result from an interaction of the specific belief issue with race, is similar to other findings (Dienstbier, 1970) that race interacts with the specific social-distance-type item used to assess prejudice. Thus, either positive or negative bias might be evidenced by the same subject toward a black stimulus individual, depending on the specific belief issue (as indicated by the present research) or the specific prejudice measure.

The specific finding from the first study which inspired the second concerned the communicator's discussion of integration. His very conservative statements ("I think it's too early for black people and white people to live together") found greater subject agreement by the white subjects of that study if the speaker were perceived as black than if white. The second study was an attempt to determine whether more radical racial information (a black-power speech) would be more consistently threatening from a black than a white communicator. In the second study, as in the first, the dependent measures indicated how similar the listener perceived his own attitudes to those expressed by the racially labeled communicator.

STUDY I

Method

Subjects. The 74 women who participated in this experiment were drawn from a women's club for newcomers to the Rochester area. Subjects were all white and generally from middle-class backgrounds; most were between 25 and 40 years old. The subjects volunteered to participate in this research in return for the experience gained from such participation and for a lecture on sex from Vincent Nowlis. Subjects were randomly assigned to the four conditions of the experiment.

Design and procedure. After being randomly assigned to one of four areas in a 360-seat auditorium, instructions were read to all four experimental groups simultaneously. Those instructions informed the subjects that they were to participate in two studies, one of the effect of "TV blurriness" on mood, the other concerning the evaluation of an (audio only) interview. The subjects were told that the person who was the subject of the silent video sequence was the same person whom they would subsequently hear interviewed. It was explained that it would be helpful to know what the interviewee looked like in order to be able to visualize him during the interview; since the video-blurriness study was to have been done anyway, the subjects were told that it was time-saving for the experimenter and the subjects to have the interviewee in the video sequence.

The subjects then filled out a Mood Adjective Check List (MACL; Nowlis, 1965), saw the one-minute video sequence which differed for each of the four groups, and filled out a second MACL. Curtains erected in the auditorium allowed subjects to see only the one TV set in front of their group. In all four video sequences the same actor was presented, dressed, and made up to appear as either white or Negro, and of apparent middle-class or lower-class status. The race change involved grease paint and a neatly trimmed Afro wig. The social class change involved only a clothing change from suit to work clothes. A between-groups factorial design of Race × Class Status was thus effected, with each subject group seeing only one of the four TV sequences. Since the purpose of the video was only to effect the race and class manipulation, all four sequences were as identical in action as possible, consisting of a dull sequence of the man entering a room and sitting down.

Since all the subjects in each group had been led to believe that they saw the same stimulus individual as had the other subject groups (the TV sequences supposedly differing only in blurriness), it was possible to present the same (audio only) tape-recorded interview of the stimulus individual to all subjects simultaneously. The interview was presented as being a training interview of a male stimulus individual by a female counseling student.

Pretests of the TV sequences on another population indicated that the actor was adequately convincing as either a black or white man. (A great deal of selecting and pretesting was required to find a man whose voice was believable as black or white.) A postexperimental check on the suspiciousness of the subjects on that aspect of experimental credibility was also conducted.

The interview lasted about 15 minutes, with topics discussed by the interviewee ranging from interpersonal relationships with family and friends to more abstract questions concerning political participation, crime, rioting, and integration. Generally, the interviewee was rather dull and negative.

Following the interview, all subjects answered scaled items which pertained to the felt similarity of their beliefs to those expressed by the stimulus individual. The purpose of the similarity items was to measure the subject's felt similarity to the beliefs expressed by the interviewee. Subjects were asked to "Please rate the man who was interviewed, Mr. Jackson, on the similarity of his views to your own." The instructions stressed that "In all cases, you will find it possible to answer the questions from your memory of your impressions of the interview you heard." Sixteen concepts that had been discussed in the interview were rated, including occupational ambition and stability, marriage fidelity, cleanliness, honesty, loyalty to friends, adequacy of encouragement for his children, educational ambition, economic wisdom, political interest, reasonableness of advocated criminal punishment, balance of views on rioting and integration, and the degree to which he drank, was religious, and superstitious. Each concept was rated on a 7-point scale with polar dimensions of "identical" and "opposite." Following the last of these scales, one scale asked the subjects to rate the overall similarity of their beliefs to those of the stimulus individual.

Following written instructions on their use, 13 modified behavioral differential items (Triandis, 1964) representing factors of formal social acceptance, friendship acceptance, social distance, and subordination were included. These were presented in the same 7-point form previously described, with polar concepts of "would" and "would not."

The behavioral differential scales were followed by a 7-item updated Ethnocentrism Scale (Adorno et al., 1950) taken from the "Other Minorities and Patriotism" section of the "suggested final form" of the Ethnocentrism Scale. Items were updated by changing references from "zoot-suiters" to "hippies," from the League of Nations to the United Nations, and from the "secret of the atomic bomb" to "military rockets" Sections pertaining to ethnic minorities were not included, since it was expected that some Jewish subjects would be participating and since not all subjects were exposed to Negroes in the experiment, possibly resulting in differential sensitization to Negro items.

The last page of the questionnaire booklet concerned suspiciousness; subjects were asked questions ranging from general ones of whether they noticed "anything special or unusual about the TV presentation" or about the interview, to specific items concerning the probable occupation of the interviewee and his race. Finally, subjects were asked to comment on the overall goals of the experiment as they saw them; a series of specific questions designed to detect suspiciousness of all critical aspects of the experiment followed.

With the completion of the questionnaires by the subjects, the experiment was over. The major aims of the research and the necessity for the race and status deceptions were then explained to the subjects in an extended debriefing.

Results

Of the 74 subjects who finished the experiment, 18 were eliminated for suspiciousness of the race manipulation, gross misidentification of the race of the stimulus individual, or the failure to follow instructions.

Although subject attrition did not vary between conditions, the loss of subjects resulted in unequal numbers of subjects in the four conditions of the experiment and the analyses of variance which were performed on the scale data were of the unweighted means type (Meyers, 1966). This design allowed the use of the data of all the subjects who were not eliminated for the reasons stated above.

Generally, as indicated in Table 1, the hypotheses concerning the effect of perceived class status on the effect of the belief material were confirmed for 13 of the 16 belief issues (in direction of effect, but not statistically significant), beliefs being rated more similar if attributed to the middle-class stimulus person. Statistical significance ($p < .05$) was found on four of those issues concerning ambition, occupational stability, loyalty to friends, and aspirations for children. A composite similarity score based on all 16 issues also showed a similar significant class effect. All three scales which indicated a reverse from hypothesis were nonsignificant.

Racial influences on belief similarity ratings were not as consistent. In fact, 11 of the 16 independent belief-topic items indicated a trend opposite to the hypothesis, with greater rated similarity in the communicator-Negro than in the communicator-white conditions. The F ratio of less than 1.0 for the composite (of all the belief-topic items) similarity score also reflected the inconsistent race effect between the similarity scales. Of the three items which indicated statistical significance based on the race variable, subjects agreed more with the white stimulus individual's views concerning aspirations for his children, but more with the black individual's views on integration and religion. Since the integration issue is of major importance for its implications for the second study of this paper, it will be emphasized here.

The portion of the interview pertaining to the stimulus person's views of prejudice was as follows:

> Stimulus individual: I think sometimes it's too early for whites and blacks to live together. I don't know—sometimes—maybe it's too early for that, for integration.
>
> Interviewer: But don't you think we should try now, anyway?
>
> Stimulus person: Sometimes it seems like we shouldn't, I don't know.

The mean rating of similarity if the stimulus person was black was 4.0, 2.7 if the person was white (based on a 7-point scale, the higher score indicating greater similarity).

Of the 13 behavioral differential items used to assess social distance of the subject toward the stimulus person, some were sensitive to race alone (such as "Would exclude from my neighborhood" and "Would accept as a close kin by marriage"), some to class alone (such as "Would treat as an equal," and "Would admire the ideas of") and one to both ("Would accept as a next-door neighbor"). Some items were sensitive to neither race nor class.

Finally, within-condition correlations of the Ethnocentrism Scale score for each subject with the subject's composite similarity ratings of the stimulus individual indicated that all such correlations were in the predicted directions, although not quite

TABLE 1

STUDY I: MEAN RATINGS OF THE COMMUNICATOR AND SIGNIFICANCE TESTS BY RACE AND STATUS CONDITION

Scale	White lower	White middle	Negro lower	Negro middle	*F*		
					Race (R)	Class (C)	R × C
Similarity scores							
Ambition	1.19	1.79	1.20	1.37	1.10	4.16*	1.25
Occupational stability	1.44	2.36	1.80	2.87	1.38	7.07*	<1
Marital faithfulness	2.50	3.00	3.10	3.62	1.22	<1	<1
Honesty	3.37	4.29	3.50	4.62	<1	3.76	<1
Loyalty	3.37	3.93	2.90	4.69	<1	5.36*	1.48
Education	1.31	1.86	1.60	1.31	<1	<1	2.21
Orderliness	2.69	2.93	1.90	3.25	<1	3.38	1.64
Money spending	2.00	3.14	3.10	2.47	<1	<1	4.84*
Aspirations for children	1.37	2.79	1.10	1.44	4.32*	5.02*	1.89
Politics	2.00	2.71	2.30	2.50	<1	1.20	<1
Criminal punishment	2.69	3.50	3.00	2.31	1.24	<1	3.65
Riots	2.69	2.64	2.40	2.06	1.67	<1	<1
Religion	1.69	1.36	2.50	3.00	8.28**	<1	<1
The supernatural	2.75	2.50	2.80	2.37	<1	<1	<1
Drinking	2.44	3.14	2.90	3.94	1.80	3.47	<1
Integration	2.69	2.64	4.40	3.56	11.17**	1.25	1.01
Overall similarity scale	1.56	2.14	2.10	2.25	2.13	2.75	<1
Composite similarity score	36.19	44.57	40.50	45.25	<1	4.60*	<1
Behavioral differential scores							
Treat as subordinate	4.69	4.00	3.50	3.75	2.20	<1	<1
Admire the ideas of	1.44	2.21	1.70	2.87	1.34	5.97*	<1
Accept as an intimate friend	1.44	2.00	1.30	2.00	<1	3.72	<1
Invite to my home	3.62	3.07	4.20	3.69	1.04	<1	<1
Treat as an equal	4.06	6.00	4.00	5.00	1.10	8.40**	<1
Elect to political office[a]	1.06	1.00	1.00	1.75	—	—	—
Exclude from my neighborhood	6.31	6.50	3.80	5.00	22.80**	2.73	1.45
Invite to my club	3.69	3.21	2.20	2.44	7.69**	<1	<1
Accept as close kin by marriage	4.94	4.43	2.80	2.94	7.58**	<1	<1
Eat with	5.87	6.14	5.90	6.19	<1	<1	<1
Accept as next-door neighbor	5.44	6.00	2.60	5.25	12.78**	10.25**	4.33*
Accept as a leader	1.12	1.21	1.50	2.00	4.33*	1.12	<1
Cooperate with in a community project	4.75	5.43	5.10	5.81	<1	1.82	<1
Behavioral differential composite score	48.44	51.21	39.60	48.69	3.08	3.36	<1

[a] No significance test from analysis of variance was possible due to a lack of homogeneity of variance; a chi-square test of the results based on the number of minimum versus nonminimum ratings yielded a chi-square of 10.34 ($df = 3$, $p = .02$).
* $p < .05$.
** $p < .01$.

significant at $p < .05$. Correlations of $-.30$, $-.43$, $-.34$, and $.06$ for communicator conditions of Negro lower class, Negro middle class, white lower class, and white middle class, respectively, indicate that while subjects' ethnocentrism tended to reduce felt similarity with the three "minority" communicators, no such effect existed for the white-middle-class communicator.

STUDY II

Method

Subjects. Subjects were 122 white male and female University of Nebraska basic psychology students who volunteered to participate in this study in order to partially fulfill an experimental participation requirement.

Design and procedure. Subjects were informed that the procedure was an attitude measurement technique and that they would hear a person discuss various topics on a tape; they were then to respond by indicating "how much you agree or disagree with the speaker, and how much you think you might like him. . . .this is really a survey of your own attitudes on the topics discussed in the tape." Before the tape was played, all subjects were instructed to read part of a professionally printed flyer which was included in the questionnaire booklet announcing a "symposium on the black in white America." It included one paragraph pertaining to the speaker whom the subjects were about to hear.

The flyer described preregistration procedures for the symposium, listed the schedule of speakers, and gave a short one paragraph biographical sketch of three of the symposium speakers. The flyer was the same for the four Race × Status manipulations except that the description of the second participant varied. That critical paragraph was circled by hand, with the accompanying notation "read this section—this speaker will be heard on tape" hand written beside the paragraph. The paragraph, in its four variations, appears below.[3]

White lower class: Thursday, 11:00. Mr. *Donald Lindzey* on "White Racism, Black Response." Mr. Lindzey *grew up in rural Arkansas where his family worked as sharecroppers. After reaching the eighth grade, he contributed to the support of his family* by working at odd jobs until joining in voter registration drives in Mississippi from 1966 to 1968. *Largely self-educated since then*, Mr. Lindzey has been *one of the few active white lecturers on the topic of black power* during the last few years. His speaking has taken him throughout the United States, including the Deep South.

White middle class: Thursday, 11:00. Mr. *Donald Lindzey* on "White Racism, Black Response." Mr. Lindzey *grew up in a suburb of Little Rock, Arkansas. His father was a lawyer. He attended the University of Michigan for two years, majoring in sociology.* He left Michigan in 1966 to participate for two years in voter registration drives in Mississippi. Mr. Lindzey has been *one of the few active white lecturers on the topic of black power* during the last few years. His speaking has taken him throughout the United States, including the Deep South.

Black lower class: Thursday, 11:00. Mr. *Howard Washington* on "White Racism, Black Response." Mr. Washington *grew up in rural Arkansas where his family worked as sharecroppers. After reaching the eighth grade in an all black school, he contributed to the support of his family* by working at odd jobs until joining in voter registration in Mississippi from 1966 to 1968. *Largely self-educated since then*, Mr. Washington has been active in lecturing on the topic of black power during the last two years. His speaking has taken him throughout the United States, including the Deep South.

Black middle class: Thursday, 11:00. Mr. *Howard Washington* on "White Racism, Black Response." Mr. Washington *grew up in a predominantly white suburb of Little Rock, Arkansas, where his family was the only black family in the neighborhood. His father was a lawyer. He attended the University of Michigan for two years, majoring in sociology.* He left Michigan in 1966 to participate for two years in voter registration drives in Mississippi. Mr. Washington has been an active lecturer on the topic of black power during the last two years. His speaking has taken him throughout the United States, including the Deep South.

That manipulation of communicator's race and class status formed, with sex of subject, a 2 × 2 × 2 factorial design.

For this second study, two different tapes were made, one attributed to a black man named Howard Washington, the second to a white man named Donald Lindzey. Several factors prompted this change from Study I. First, it is extremely difficult to make a single sound tape which is equally convincing as that of a white or black speaker. Second, the introduction of vocal differences associated with race would increase the power of the race manipulation. It was felt that these considerations more than balanced the advantage of the clean experimental interpretation allowed by the single-tape manipulation. The procedure necessitated, of course, the running of two separate groups of subjects, each hearing only one of those two critical audiotapes. Attempts to eliminate any selection bias between those two groups included scheduling both sessions at a time in the late afternoon when classes were few, and running the two sessions with the second following immediately after the first. Within each group, subjects of both sexes were randomly assigned to the status conditions. Strenuous attempts were made to equate the two audiotapes on overall quality and emphasis. Specifically, although the verbal message was quite controversial, both speakers maintained a calm and paced delivery; this seemed appropriate in view of the fact that the speech was supposedly delivered and recorded at a university symposium.

The opinions on which subject agreement was assessed had all been pretested (the previous semester) in written form for agreement with an independent sample of basic psychology students. It was thus ascertained that a fairly wide rage of statement acceptability was achieved. (Agreement ranged from 4.95 to 2.28, respectively, on a 7-point belief similarity scale from "identical" to "opposite" for items of "Black people do not want to move in with white people" and "Half of every tax dollar from white people should go to support black education, black social services, and black culture.")

The tape recording lasted approximately five minutes, though subjects were left with the impression that it was only the first part of the communicator's speech which was played for them. Unlike Study I, the topics of the speech were restricted to issues of the black minority in a white culture. The range of topics can be seen by the item headings on the first seven items of Table 2.

---

[3] Italics are presented here only to accentuate critical differences between paragraphs and were *not* included in the study manipulations.

After the tape was played, subjects were instructed to turn to the page in their questionnaire booklets which contained the scaled items related to the speech. Instructions for the opinion similarity rating told subjects to "rate the similarity of your own views to his views on the following topics. Base your ratings on what you heard Mr. Washington [or Mr. Lindzey] say on the tape recording." Ratings were made on the same Lykert-type 7-point scales with polar adjectives of "identical" and "opposite" which were used in the pretest for Study II and in Study I. The last of those scales asked subjects for a rating of the general similarity of their values to those of the speaker.

Finally, five social distance and behavioral differential scales similar to some of those used in Study I were employed, with instructions to the subjects to "assume that you had the opportunity to interact with him in the manner suggested by the question" in responding to those items. Those scales form the last five items of Table 2.

After filling out that form, subjects were given a questionnaire meant to assess suspiciousness concerning any deception in the study. The form explained that subjects often "form their own explanations or hypotheses about what the experimenter is 'really' after," and asked a series of open-ended questions about what suspicions the subject had formed, how certain he was of those suspicions, etc.

Subjects were debriefed concerning the true nature and purpose of the experiment.

Results

Of the 122 subjects who took part in the experiment, two were eliminated for suspiciousness that the tape recordings were not real. Subjects were randomly eliminated from seven of the eight cells of the analysis of variance design to achieve an equal number of subjects per cell, bringing the number of subjects used in the analyses to 88 (there being 11 subjects in the smallest cell). It was felt that the rather large number of subjects still in the design would provide an accurate assessment of the independent variable effects without necessitating the unweighted-means-type analysis used in Study I.

Table 2 presents the major results of Study II. Results of communicator variables by subject sex are not included in that table since sex was never a significant

TABLE 2

STUDY II: MEAN RATINGS OF THE COMMUNICATOR AND SIGNIFICANCE TESTS BY RACE AND STATUS CONDITION

Scale	White lower	White middle	Negro lower	Negro middle	*F*		
					Race (R)	Class (C)	R × C
			Similarity scores				
White's desire for integration	4.27	4.14	3.41	4.41	<1	1.81	3.14
Black's desire for integration	4.14	4.91	3.91	4.09	2.34	1.94	<1
White racism in American education	4.46	4.77	3.73	4.36	2.36	1.66	<1
Black's right to grab what's rightfully his	3.18	4.00	3.14	2.86	2.47	<1	2.10
Justice for blacks in America	4.82	5.36	4.18	4.41	4.75*	1.12	<1
Rioting and burning by blacks	3.23	3.55	2.00	1.86	15.50**	<1	<1
White support of blacks through taxation	3.36	3.91	3.59	3.05	<1	<1	2.18
General similarity of your values	3.55	3.95	2.82	3.23	4.83*	1.53	<1
			Social distance scores				
Similarity to you	3.45	3.68	2.18	2.68	13.23**	1.35	<1
Liking if met	4.09	4.55	3.00	3.59	7.64**	1.99	<1
Would elect to political office	3.36	3.77	2.32	2.59	8.55**	<1	<1
Would accept as close kin by marriage	4.45	4.77	2.55	3.64	13.32**	2.85	<1
Would accept as a next-door neighbor	5.41	5.44	3.82	5.05	6.45*	2.62	2.25

* $p < .05$.
** $p < .01$.

interactive variable with race or class. On 12 of the 13 similarity and social distance variables, women subjects tended to be more accepting and to see their beliefs as more similar to the communicator's reversing only on the issue of election to political office. The only statistically significant effects related to sex were for the liking-if-met item (mean liking by women = 4.20 compared to 3.41 for men, $F = 4.61$, $df = 1/80$, $p < .05$) and the accept-as-kin item (mean acceptance by women = 4.30 compared to 3.41 for men, $F = 4.51$, $df = 1/80$, $p < .05$).

Unlike Study I, in which status seemed the most consistently reliable determinant of agreement variance (the communication in Study I did not largely concern race-relevant issues), the results of Study II indicated that communicator's race was the single most powerful influence on variance for the similarity and social distance scales. That finding is indicated strongly by greater agreement with and liking for the white communicator in every one of the 13 scales used in Study II. This race effect is significant for three of the eight similarity scales (including the general similarity item), and for all five of the social-distance-type measures.

Although the direction of the status differences were consistent with those of Study I, with the middle-class communicator receiving more favorable ratings on 12 of the 13 dependent measure scales, the differences are not statistically significant on any measure.

Finally, there are no statistically significant Race × Class interactions.

DISCUSSION OF THE DATA

Between Study I and Study II, differences existed (*a*) in the manner in which independent variables of race and status were manipulated, (*b*) in the content and style of communication delivery, (*c*) in some of the dependent measure scales, and (*d*) in the characteristics of the subject populations used. Yet, if the findings of the two studies are broadly considered, it is possible to compare and contrast those findings.

Across both studies, the manipulation of social class status resulted in subjects rating their beliefs as less similar to the lower-class communicator, compared to the middle-class communicator. The topics discussed in the two scripts of the two studies did not, of course, exhaust possible and likely topic areas, and in Study II, while the script was purposefully made race relevant, no such effort was made for relevancy to social class. Therefore, despite the consistency of the class findings, it is expected that one could find topics for such research presentations with which even middle-class subjects would find more agreement with the lower-class communicator.

Within Study I, communicator race and communication topic interacted, so that some issues induced more agreement with the white communicator, while others resulted in the black receiving more agreement. The complexity of the interactions between communicator characteristics and communication content is, therefore, more evident with respect to race than class variables. By restricting the range of the topics used in Study II to a single class of race-relevant issues, it was predicted that the irregular race-topic interaction found in Study I would give way to a consistent unidirectional race effect; this was realized in Study II as the consistent tendency to agree more with (and to like more) the white communicator.

Taken together, the two studies illustrate that expectations which subjects have about class are different from those about race, and that a communication which is relevant to class and/or race, and which is presented in a fashion which permits interpretation latitude, interacts in unique ways with those variables. It is apparent from Study I that these interactions between information content and communicator-characteristic variables are mediated by subject characteristics associated with expectencies, sets, or stereotypes relevant to those communicator characteristics; this relationship was illustrated by the pattern of correlations between subject ethnocentrism and felt similarity toward the communicator. It is also evident, particularly from the results of Study I, that no simple formulation such as suggested by congruity theory (Osgood & Tannenbaum, 1955) based only on the valences of the communicator and the communication could account for the complex ways in which communication and communicator characteristics interact. As illustrated by the behavioral differential and social distance measures of both studies of this paper (as well as by Dienstbier, 1970; Triandis, 1961; Triandis & Davis, 1965), the degree to which prejudice would be indicated as due to status or race would depend on the dependent measures used to operationalize that construct. It is difficult to avoid the suspicion that some circularity has crept into the (reviewed) belief theory literature with respect to the choice of dependent measures. Prejudice has often been operationally defined by some dependent measures which have proven to be more sensitive to the belief manipulation than the race manipulation. To suggest, as did Mezei (1971), that a more race-sensitive measure is unfair since it would be related to and confounded with perceived social pressure is on the one hand admitting the complexity of prejudice, while on the other denying the importance of that complexity since it does not fit the simpler belief theory.

IMPLICATIONS FOR A REVISED BELIEF THEORY

The belief theory of prejudice suggests that when we perceive an individual to be black, then we make assumptions about his different beliefs—it is these assumptions which account for our negative feelings and our discriminatory actions toward that individual. Not spelled out in this formulation, but seemingly obvious, is the stipulation that the belief areas about which assumptions are made will be relevant to the individual's race, or our stereotype of it. If we see a black person we will not necessarily assume that he is a communist, or an anarchist, or that be believes anecdotal evidence is more valuable than experimental psychology in providing reliable information. Rather, the belief dimensions upon which we might base our discrimination would be relevant to our stereotypes about individuals of that race, and would, therefore, be solidly in the areas about which we anticipate nonconsensus with those indiviuals. It makes little sense, then, to attempt to demonstrate that racial differences of stimulus individuals have little power to influence us when those stimulus individuals are equated on a host of irrelevant beliefs (as has been done so often in the cited belief theory literature). This exercise can only show us that race is not interactive with *those* belief issues, and that subjects are reluctant to respond negatively toward a racially labeled stimulus person without some justification or that they have no racial prejudice to begin with (or are reluctant to show it), or that they are making unwarranted assumptions about what they should respond to, or that they do not believe the experimenter.

It is apparent that the theoretical and experimental considerations are (or should be) more complex than suggested by the unidirectional conceptualizations of the belief theory researchers. The attempt to account for complex constructs like prejudice on the basis of a single underlying principle cannot succeed. This pattern of theory construction represents a way of conceptualizing human behavior whose roots lie plainly in an oversimplistic view of man.

It may indeed be true (and quite apparently is so) that one's assumption that another individual will not validate one's belief systems will dispose one negatively toward that person. But it is equally true, as evidenced by the present research, that when the belief information is relevant to racial issues, those assumptions of dissimilar beliefs are not always simply and quickly allayed by the presentation of belief information. It appears from this data that assumptions about belief differences are caused by prejudice (read antipathy or negative affect) just as it is *also* apparent from the belief theory research that prejudice is caused by assumptions about belief differences. In the first case the negative affect may exist for any number of reasons, as indicated by Allport (1954) and discussed in the introduction to this paper. We are not left with a choice between either race (i.e., responses to race not based on belief issues) or belief to account for prejudice. There may indeed be times when a greater reliance on one or the other of those constructs may provide us greater success in the prediction, understanding, and control of certain behaviors, but these are concepts so different in structure that they can neither be logically compared for relative strength nor can they be heuristically seen as hierarchically arranged. They hold, instead, relationships with each other most fruitfully seen as mutually causal or interactive. To lose sight of this mutual causation is to oversimplify the area of prejudice, with the ultimate result that our laboratory studies will have less and less relevance for the real world.

REFERENCES

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Wiley, 1950.

Allport, G. W. *The nature of prejudice.* Garden City: Doubleday, 1954.

Blake, R., & Dennis, W. The development of stereotypes concerning the Negro. *Journal of Abnormal and Social Psychology,* 1943, **38**, 525–531.

Byrne, D., & Erwin, C. R. Attraction toward a Negro stranger as a function of prejudice, attitude similarity, and the stranger's evaluation of the subject. *Human Relations,* 1969, **22**, 397–404.

Byrne, D., & Mc Graw, C. Interpersonal attraction toward Negroes. *Human Relations,* 1964, **17**, 201–213.

Byrne, D., & Wong, T. J. Racial prejudice, interpersonal attraction, and assumed dissimilarity of attitudes. *Journal of Abnormal and Social Psychology,* 1962, **65**, 246–253.

Dienstbier, R. A. Positive and negative prejudice: Interactions of prejudice with race and social desirability. *Journal of Personality,* 1970, **38**, 198–215.

Dollard, J. *Cast and class in a southern town.* New Haven: Yale, 1937.

Hendrick, C., Bixenstine, E., & Hawkins, G. Race versus belief similarity as determinants of attraction: A search for a fair test. *Journal of Personality and Social Psychology,* 1971, **17**, 250–258.

Insko, C. A. *Theories of attitude change.* New York: Appleton-Century-Crofts, 1967.

Insko, C. A., & Robinson, J. E. Belief similarity versus race as determinants of reactions to Negroes by southern white adolescents: A further test of Rokeach's theory. *Journal of Personality and Social Psychology,* 1967, **7**, 216–221.

Kirscht, J. P., & Dillehay, R. C. *Dimensions of authoritarianism: A review of research and theory.* Lexington: University of Kentucky, 1967.

Meyers, J. L. *Fundamentals of experimental design.* Boston: Allyn-Bacon, 1966.

Mezei, L. Perceived social pressure as an explanation of shifts in the relative influence of race and belief on prejudice across social interactions. *Journal of Personality and Social Psychology,* 1971, **19**, 69–81.

Nowlis, V. Research with the Mood Adjective Check List. In S. S. Tomkins & C. E. Izard (Eds.), *Affect, cognition and personality.* New York: Springer, 1965.

Osgood, C. E., & Tannenbaum, P. H. Congruity in the prediction of attitude change. *Psychological Review,* 1955, **62**, 42–55.

Page, M. M. Postexperimental assessment of awareness in attitude conditioning. *Educational and Psychological Measurement,* 1971, **31**, 891–906.

Robinson, J. E., & Insko, C. A. Attributed belief similarity-dissimilarity versus race as determinants of prejudice: A further test of Rokeach's theory. *Journal of Experimental Research in Personality,* 1969, **4**, 72–77.

Rokeach, M., & Mezei, L. Race and shared belief as factors in social choice. *Science,* 1966, **151**, 167–172.

Rokeach, M., Smith, P. W., & Evans, R. I. Two kinds of prejudice or one? In M. Rokeach (Ed.), *The open and closed mind.* New York: Basic Books, 1960.

Smith, C. R., Williams, L., & Willis, R. H. Race, sex and belief as determinants of friendship acceptance. *Journal of Personality and Social Psychology,* 1967, **5**, 127–137.

Stein, D. D. The influence of belief systems on interpersonal preference: A validation study of Rokeach's theory of prejudice. *Psychological Monographs,* 1966, **80** (8, Whole No. 616).

Stein, D. D., Hardyck, J. A., & Smith, M. B. Race and belief, an open and shut case. *Journal of Personality and Social Psychology,* 1965, **1**, 281–289.

Triandis, H. C. A note on Rokeach's theory of prejudice. *Journal of Abnormal and Social Psychology,* 1961, **62**, 184–186.

Triandis, H. C. Exploratory factor analysis of the behavioral component of social attitudes. *Journal of Abnormal and Social Psychology,* 1964, **68**, 420–430.

Triandis, H. C., & Davis, E. E. Race and belief as determinants of behavioral intentions. *Journal of Personality and Social Psychology,* 1965, **2**, 715–726.

(Received October 8, 1971)

Psychological Review
1972, Vol. 79, No. 2, 161–171

A STIMULUS-SAMPLING MODEL OF THE PARTIAL REINFORCEMENT EFFECT

RONALD L. KOTESKEY[1]

Asbury College

A stimulus-sampling model of E. J. Capaldi's sequential effects hypothesis to account for the increased resistance to extinction following partial reinforcement was presented in axiomatic form. The model was shown to be capable of accounting for the major relevant variables which produce the increased resistance to extinction, and it was shown to not predict the increased resistance to extinction where it does not occur. Several limitations of the model were discussed as were the model's relationships to other theories.

Stimulus-sampling theory has never been able to adequately account for the increased resistance to extinction following partial reinforcement during acquisition, commonly called the partial reinforcement effect. Capaldi's (1966) hypothesis of sequential effects has been able to account for the partial reinforcement effect; however, a statement of the theory as a stimulus-sampling model has several potential advantages. Once a model has been stated as a set of axioms in stimulus-sampling terms, a mathematical or computer simulation model of the theory can be constructed and this process of building such a model forces every assumption to be stated explicitly. After the model has been constructed, it can be used to make precise predictions and these predictions can be checked and shown to follow logically from the axioms. Since quantitative predictions can often be made with such a model, one is in a better position to evaluate how well the theory accounts for the data. If the predictions are wrong, the model can often be modified and new predictions made to see if the modification has enabled the theory to account for the data.

Koteskey and Hendrix (1971) proposed such a stimulus-sampling model of Capaldi's (1966) sequential interpretation of the partial reinforcement effect. Although the model was briefly presented there with an experiment to test one of its predictions, space limitations restricted discussion of the model at any length. This article discusses the model, how it can account for the major variables, its limitations, and its relationships to other theories.

THE MODEL

The assumptions presented in Koteskey and Hendrix (1971) are stated more formally here as axioms:

Representation Axioms

1. There are N_T stimulus elements.
2. There are two responses labeled A_1 (respond within the time limit) and A_2 (fail to respond within the time limit). One of these occurs on each trial.
3. There are two reinforcing events E_1 (reinforcement which reinforces A_1) and E_2 (nonreinforcement which reinforces A_2). One of these occurs on each trial.

Conditioning Axioms

1. Each element is conditioned to one and only one of the responses.
2. If an element is sampled on a trial, it becomes conditioned with probability c to the response that is reinforced on that trial; with probability $1 - c$ it remains conditioned as before. If the sampled element is already conditioned to the reinforced response, it remains so.
3. Stimulus elements that are not sampled on a trial do not change their states of conditioning on that trial.
4. At the beginning of training, all N_T elements are conditioned to A_2.

---

[1] Requests for reprints should be sent to Ronald L. Koteskey, Department of Psychology, Asbury College, Wilmore, Kentucky 40390.

Stimulus-Sampling Axioms

1. During each trial, a sample of elements is taken.
2. Following each reinforcement, the same subset of elements, labeled N_0, is sampled.
3. Following each successive nonreinforcement, additional elements are included in the sample, the number of additional elements being a constant proportion, p, of the remaining elements. These additional subsets of elements are labeled N_1, N_2, N_3, $\cdots$.

Response Axiom

1. The probability of response A_j on a particular trial is given by the proportion of the elements sampled on that trial which are conditioned to A_j.

Stimulus-Sampling Axiom 3 has been made more specific than in Koteskey and Hendrix (1971) as shown in Figure 1; rather than arbitrarily assigning values to N_1, N_2, N_3, $\cdots$, the size of each of these sets of stimulus elements is now assumed to be a decreasing function so that as $i \rightarrow \infty$, the number of elements in N_i approaches zero. That is, as Capaldi (1966) points out, less stimulus modification occurs on each successive trial as length of nonreinforced run of trials (N-length) increases. The basis for this assumption is evidence cited by Capaldi (1966) in which experiments on N-length, number of different N-lengths, and number of occurrences of each N-length show the assumption to be plausible. This assumption also makes it possible to write a computer program simulating the model which will give predictions for the long N-lengths found in extinction without arbitrarily picking values of N_1, N_2, N_3, $\cdots$, which fit that particular set of results.

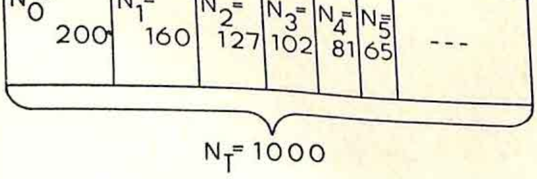


FIG. 1. Schematic representation of the stimulus-sampling process showing the parameters used. (Each successive N_i is .20 of the remaining N_T.)

The computer program makes it possible to predict resistance to extinction for any schedule of reinforcement, since it computes the proportion of the stimuli sampled on any trial that are conditioned to responding and the number of each of the N_0, N_1, N_2, N_3, $\cdots$, elements that are conditioned to responding. To show that the model can account for a great deal of data without manipulating the parameters at all, every time the computer program was used here, the following parameters were used: $N_0 = 200$, $N_T = 1{,}000$, $c = .15$, and $p = .20$. All groups were run for 100 trials unless otherwise stated. Figure 1 is a schematic representation of the stimulus-sampling process with the new sampling axiom and the parameters used throughout this article. Note that each successive N_i is a constant proportion of the remaining unsampled elements in N_T. After each successive nonreinforcement, the elements in the corresponding N_i are added to the sample.

In some cases the output from the program will be presented trial by trial, in others only the total number of elements conditioned at the end of acquisition will be presented, since resistance to extinction (persistance) is directly related to this total number of elements conditioned when extinction begins. To fit the data from any specific experiment, the parameters would have to be estimated from that data, since at least the sizes of the N_i would differ in different experimental situations and at least the sizes of p and c would differ with different species.

MAJOR VARIABLES

Since the curves presented in the following figures were all made using the same set of parameters, it should be kept in mind that it is the general shape of the curves that is important, not their exact details which change with different parameters. These parameters were chosen because they show the single alternation pattern emerging within the first 20 trials which makes a reasonable-sized figure. Other parameters would change the point at which this pattern would emerge, the peak of the inverted U found in percentage of reinforcement, etc.

Increased Resistance to Extinction Occurs and Is Accounted For

The partial reinforcement effect has been shown to occur in the following situations, and the magnitude of the effect has been shown to be a function of some of these variables. The first three of these have been confounded in a number of studies.

Percentage of reinforcement. Lewis (1960) concluded in his review of partial reinforcement that there was an inverted-U relationship between the magnitude of the partial reinforcement effect and percentage of reinforcement; 100% and 0% reinforcement show little resistance to extinction, and the peak of the U varied with level of acquisition. Weinstock (1954, 1958) found that with groups between 16% and 100% reinforcement and one trial per day, increasing percentage of reinforcement resulted in a decreased partial reinforcement effect. Coughlin (1970) using eight trials per day found increasing resistance to extinction down to 25% reinforcement but decreasing resistance to extinction with 12.5% reinforcement. Koteskey and Stettner (1968) using nine trials per day found that some subjects on a 33% reinforcement schedule refused to run after about 200 acquisition trials and had a greatly reduced resistance to extinction. So, 10 years later, existing data continue to support Lewis's (1960) conclusion that other things being equal there is an inverted-U relationship between percentage of reinforcement and resistance to extinction.

In the model, as percentage of reinforcement decreases from 100% to at least 50% in an essentially random schedule, there are longer runs of consecutive nonreinforcements as well as more of these long runs of nonreinforcements. Therefore, more elements (such as N_2, N_3, N_4, $\cdots$) are included in the sample more often so that more of them become conditioned to running. However, at some point (usually below 50% reinforcement) there will be too few reinforcements for enough conditioning to take place so that further decreasing the percentage of reinforcement will result in less resistance to extinction. The computer simulation of the model was run with percentages of reinforcement varying from 20% to 100%. The reinforcement schedule was random subject to the restriction that Trials 1 and 100 be reinforced and no run of consecutive nonreinforcements be longer than five. The total number of elements conditioned at the end of the 100 trials is shown in Figure 2. The drop in the number of elements conditioned in the 50% group was somewhat surprising, but it was found that the 50% schedule, although random, contained no run of consecutive nonreinforcements longer than three, while the 60% group contained runs of nonreinforcement of Length 4 and the 40% group contained runs of nonreinforcement of Lengths 4 and 5. Thus, the dip in the curve at 50% was a result of the particular random schedule that happened to be generated for it.

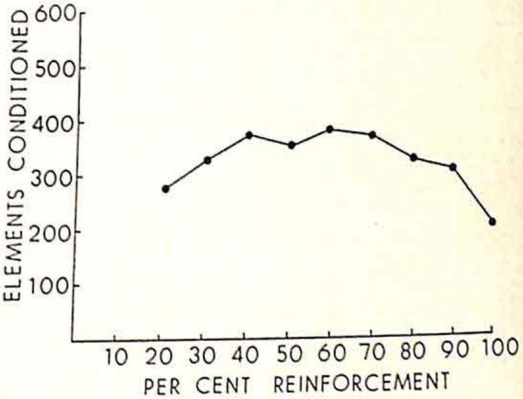


FIG. 2. Total number of elements conditioned at the end of 100 trials as a function of percentage of reinforcement.

Number of nonreinforcements. Jernstedt (1971) found that increasing the number of nonreinforcements increased resistance to extinction, and Lawrence and Festinger (1962) conducted an experiment with a factorial design to separate the effects of percentage of reinforcement and number of nonreinforcements. In previous studies these two variables had been confounded and this study showed that the effect was due to number of nonreinforcements, not percentage of nonreinforcement. In their experiment, when percentage of reinforcement was held constant, the number of nonreinforcements was increased by giving more trials. According to the model, this results in about the same number of elements being sampled, but they are sampled *more* times during acquisition, thus more of them become conditioned resulting in greater resistance to extinction. When number of nonreinforcemnets is held constant, the percentage of reinforcement is decreased by decreasing the number of trials given, so that although more elements are sampled and have an opportunity to become conditioned, this is offset by the fact that they have *fewer* opportunities to become conditioned and the net result is no increase in resistance to extinction. In Figure 2 percentage of reinforcement is confounded with number of nonreinforcements. Therefore, the abscissa may also be read as number of reinforcements to show that, up to a point, the effect of increasing the number of nonreinforcements results in increased resistance to extinction.

N-length. Capaldi (1964, 1966), Capaldi, Hart, and Stanley (1963), Capaldi and Kassover (1970), Capaldi and Minkoff (1969), Capaldi and Spivey (1963), and Gonzalez and Bitterman (1964) have all shown that the length of run of successive nonreinforcements (N-length), number of different N-lengths, and number of occurrences of each N-length are all variables relevant to the magnitude of the increased resistance to extinction. These were also confounded with percentage of reinforcement in earlier studies. In terms of the model, as N-length increases, more elements are included in the sample to be conditioned and finally results in increased resistance to extinction. The computer program of the model was run with varying N-lengths as shown in Figure 3. The group shown as *N-length 5* was given four runs of N-length 5, three of N-length 4, three of N-length 3, three of N-length 2, and three of N-length 1; the group shown as *N-length 4* was given five runs of N-length 4, five of N-length 3, five of N-length 2, and five of N-length 1; the group shown as *N-length 3* was given eight runs of N-length 3, nine of N-length 2, and eight of N-length 1; the group shown as *N-length 2* was given 17 runs of N-length 2 and 16 runs of N-length 1; the group shown as *N-length 1* was given 50 runs of N-length 1 (single alternation); and the group shown as *N-length 0* was given continuous reinforcement. Figure 3 shows the total number of elements conditioned at the end of the 100 trials in each of the above schedules, and it is obvious that as N-length and number of different N-lengths are increased, resistance to extinction is increased.

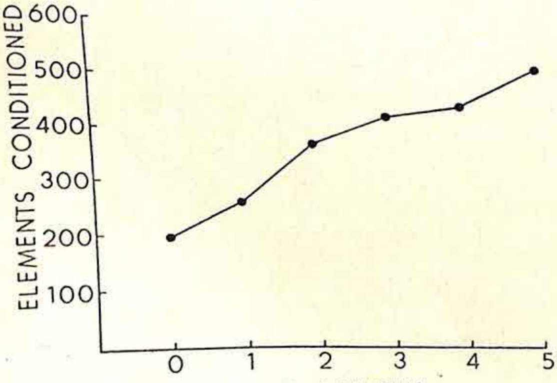


FIG. 3. Total number of elements conditioned at the end of 100 trials as a function of N-length.

Level of acquisition. Bacon (1962), Erlebacher (1963), Hill and Spear (1963), and Wilson (1964) found that increasing the level of acquisition increased resistance to extinction, at least up to about 300 trials. Since increasing the level of acquisition results in the elements being present more times to be conditioned, more of them do become conditioned and this is shown by an increased resistance to extinction. Figure 4 shows the effect of an increasing number of acquisition trials on the number of elements conditioned. The curves represent the number of elements conditioned at the end of 25, 50, 75, and 100 trials for a 50% random group and for a 100% group. In the partially reinforced group, the number of elements conditioned increased at least up to 75 trials, while the 100% group has almost reached asymptote at 25 trials with these parameters.

Single alternation: Few trials. Related to level of acquisition are the findings of Capaldi and Hart (1962) and Capaldi and Wargo (1963), whose studies showed a greater partial reinforcement effect with

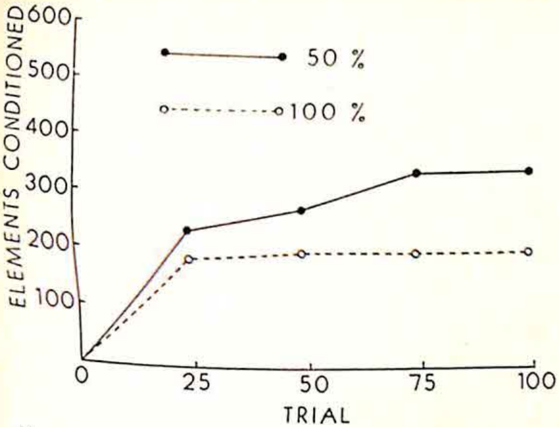


FIG. 4. Number of elements conditioned as a function of level of acquisition in a 50% random and a 100% reinforcement schedule.

single-alternation groups than with "random" groups using a small number of acquisition trials. Their "random" groups were given 2 or 3 runs of N-length 4 and the single-alternation groups were given 8 or 12 runs of N-length 1. This would result in only a few of the N_1, N_2, N_3, and N_4 elements being conditioned in the "random" groups (only present 2 or 3 times when reinforcement was given), but in a large number of the N_0 and N_1 elements being conditioned (present 8 or 12 times when reinforcement was given) in the single-alternation group, resulting in more resistance to extinction.

Double alternation. Bloom and Capaldi (1961), Habu and Ono (1969), and Koteskey and Hendrix (1971) found increased resistance to extinction in their double-alternation groups. This would be explained by the model because at the end of acquisition with a very large number of trials, nearly all of the N_2 elements would be conditioned (since they were never followed by a nonreinforcement) along with many of the N_0 and N_1 elements. Thus, with this larger number of elements conditioned, the increased resistance to extinction would be expected. In the computer run of the model with the parameters used here, the double-alternation group had 327 elements conditioned while the continuous group had only 199 conditioned.

Block of continuously reinforced trials. Theios (1962) showed that giving a block of continuously reinforced trials to a partially reinforced group did not significantly decrease resistance to extinction. Of course, this model would say that once elements in N_1, N_2, N_3, $\cdots$, were conditioned during acquisition, they would not be sampled during the block of continuous reinforcement, so they would remain conditioned and produce the increased resistance to extinction. Sutherland, Mackintosh, and Wolfe (1965) found that if the partially reinforced block of trials was given early in training and then continuous reinforcement given, this group had a larger partial reinforcement effect than a group given continuous reinforcement first, then partial. The model would say that the reason for the increased resistance to extinction is that when continuous reinforcement is given later, more of the N_0 elements are conditioned when extinction starts, producing the increased resistance to extinction. Support is given for this interpretation in that Theios and McGinnis (1967) pointed out that the Sutherland et al. (1965) groups differed in response speed at the end of acquisition and their results may be due to this factor. Campbell, Knouse, and Wroten (1970) and Koteskey and Hendrix (1971) also found that continuous following partial reinforcement resulted in increased resistance to extinction. This is because even though the N_1, N_2, N_3, $\cdots$, elements are not affected, more and more N_0 are becoming conditioned during the continuous reinforcement. For instance, the double-alternation group mentioned above which has 327 elements conditioned at the end of 100 trials would have 409 elements conditioned if the 100 trials were followed by a block of 25 continuously reinforced trials.

Miniature experiments. The partial reinforcement effect has been found in some of the miniature experiments of Amsel, Hug, and Surridge (1968), Bowen and McCain (1967), Capaldi and Waters (1970), McCain (1965a, 1965b, 1965c, 1966, 1968), and McCain and Brown (1967). In these miniature experiments acquisition consists of a very small number of trials, sometimes as few as two, a nonreinforcement and a reinforcement. The model would sometimes predict a partial reinforcement effect in these experiments because after the nonreinforced trial, elements from N_1 would be present to be conditioned on the following reinforced trial; however, whether or not the effect would be predicted would depend on the parameters estimated in a particular experiment. For instance, with the parameters used here, a two-trial experiment would have 53 elements conditioned in the partially reinforced group and 55 elements conditioned in the continuously reinforced group. Thus, the model would not predict that the increased resistance to extinction would be found in every study, such as Surridge, Rashotte, and Amsel (1967).

Increased Resistance to Extinction Does Not Occur

Some situations have been discovered in which increased resistance to extinction is not found, is greatly reduced, or is generalized. The model not only predicts the increase in resistance to extinction when it should be found, but predicts that it should not be found in situations where it is not found.

Single alternation: Many trials. If the reinforcement schedule is a single-alternation schedule and is continued for a large number of trials, a greatly reduced partial reinforcement effect, if any, is found. Bloom and Capaldi (1961), Campbell et al. (1970), Koteskey and Hendrix (1971), and Tyler, Wortz, and Bitterman (1953) found a reduced partial reinforcement effect in single-alternation groups given a large number of trials as compared to randomly reinforced or double-alternation groups. Also, the investigators in several of these articles noted the following pattern in speed of response: (*a*) early in acquisition the subjects ran faster following a reinforced trial than following a nonreinforced trial; (*b*) at intermediate stages of acquisition this difference disappeared; and (*c*) later in acquisition the opposite was true—they ran faster following a nonreinforced trial.

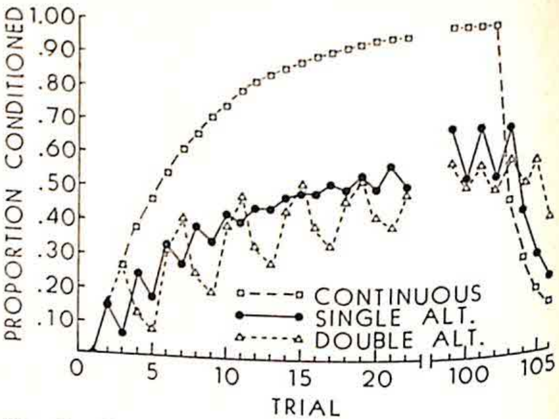


FIG. 5. Proportion of elements in the sample conditioned at the beginning of each trial. (Odd-numbered trials are reinforced in single alternation, Trials 1, 2, · · ·, 5, 6, · · ·, 9, 10, etc., are reinforced in double alternation, and all trials are reinforced in the continuous group. Extinction starts on Trial 102.)

In running the model on the computer, changing parameters will produce greater, less, or equal resistance to extinction in the single-alternation group as compared to a continuously reinforced control group, but regardless of the amount of resistance to extinction in the single-alternation group, the model gives the above response pattern as shown in Figure 5 which *is* found experimentally. The reasons this pattern emerges are the following: (*a*) On the first trial (reinforced), some of the elements in N_0 are conditioned; on the second trial (nonreinforced), the elements from N_0 are sampled and the proportion conditioned determine running speed, but since it is nonreinforced, some of the elements in N_0 become unconditioned; on the third trial (reinforced), the elements from N_0 and N_1 are sampled, and since some of the conditioned elements in N_0 became unconditioned on the previous trial *and* all of the N_1 elements are unconditioned, this lower proportion of conditioned elements results in a slower running speed; on the fourth trial (nonreinforced), the elements from N_0 are sampled, and since a larger proportion are conditioned, a faster running speed is the result. The above pattern continues over the first stages of acquisition. (*b*) Since none of the trials on which the elements in N_1 are sampled are followed by nonreinforcement (every nonreinforcement is followed by a reinforcement), none of the elements in N_1 reenter the unconditioned state during acquisition so at intermediate stages of acquisition the proportion of elements conditioned after both reinforcement and nonreinforcement is roughly equal resulting in no difference in response speed. (*c*) As more and more of the N_1

elements become conditioned as acquisition continues (all N_1 elements become conditioned at asymptote), it finally reaches the point where the proportion of conditioned elements in the sample is greater following a nonreinforcement than following a reinforcement, the opposite of early in acquisition.

The higher proportion of elements conditioned on trials following reinforcement in the double-alternation as well as the single-alternation group early in acquisition is consistent with Bloom (1967) and Bloom and Capaldi (1961), in which running speed was faster on trials following reinforcement early in acquisition whether the schedule was regular or irregular. The model generally predicts a faster running speed on trials following reinforcement, since elements in N_0 are sampled and more of them are conditioned than any other N_i early in acquisition. Thus, the model can even account for specific patterns of response during acquisition.

No nonreinforcement–reinforcement transitions. Grosslight and Radlow (1956, 1957) and Koteskey and Stettner (1968) found no partial reinforcement effect in groups in which no nonreinforcements were followed by a reinforcement within a day's block of trials when several trials were given in a day. Since the animals are fed at the end of a day's block of trials, this has nearly the same effect on sampling as a reinforcement, so that on the first trial of the next day, they sample only the N_0 elements and very few, if any, of the N_1, N_2, N_3, $\cdots$, elements. Thus, few of the elements added to the sample after nonreinforcement ever become conditioned, and this group shows little or no partial reinforcement effect. In fact, this group may show slightly less resistance to extinction than the continuous reinforcement group, since some of the N_0 elements become unconditioned on the nonreinforced trials, and this is the trend in all three studies cited above, although none of the differences were significant.

The number of nonreinforcement–reinforcement transitions has even been shown to be crucial in determining the amount of increased resistance to extinction by Koteskey and Stettner (1968), Spivey (1967), and Spivey and Hess (1968). Nonreinforcement–reinforcement transitions alone are not enough to determine the size of the increase in resistance to extinction but some are necessary for it to occur.

Within-subject experiments. In the within-subject experiments by Amsel (1967), Brown and Logan (1965), Galbraith, Rashotte, and Amsel (1968), Pavlik and Carlton (1965), Rashotte and Amsel (1968), and Rashotte, Ross, and Amsel (1968), a generalized partial reinforcement effect has usually been found; that is, the response in the partially reinforced alley was no more resistant to extinction than the one in the continuously reinforced alley, although they are both more resistant to extinction than subjects given only continuous reinforcement. According to the model, this is because some of the trials in the continuously reinforced alley follow a nonreinforcement in the partially reinforced alley, and since they follow a nonreinforcement or series of nonreinforcements, N_1, N_2, N_3, $\cdots$, are present in the continuously reinforced alley and result in increased resistance to extinction in that alley as well as the partially reinforced alley. Even data which at first glance seem to be incompatible with this interpretation, such as those presented by Rudy, Homzie, Cox, Graeber, and Carter (1970), are followed by the authors' conclusion that a sequential interpretation is needed to account for it.

Intertrial reinforcement: Small number of trials. Subjects given intertrial reinforcement after each series of nonreinforced trials (before a reinforced one) show no partial reinforcement effect with a relatively small number of trials as shown by Capaldi et al. (1963) and Capaldi and Spivey (1963). Again the model would say that this is because the reinforcement in the goal box affects the sampling in nearly the same way as does a reinforced trial, so that only the elements in N_0 and possibly a few N_1, N_2, N_3, $\cdots$, elements are present to become conditioned. Thus, all the elements usually present following nonreinforcement can never become conditioned, so little or no partial reinforcement effect results.

Increased Resistance Occurs but Is Not Specifically Accounted For

As with most models there are certain phenomena which occur that are not specifically accounted for by the model. In this section, some examples are selected to show that although these types of phenomena exist for this model, they are not in conflict with it; it is only that the model as stated does not specifically include them.

Intertrial reinforcement: Large number of trials. Berg, Connelly, and Duerfeldt (1969), Black and Spence (1965), and Spence, Platt, and Matsumoto (1965) found a partial reinforcement effect when a large number of acquisition trials were used even with intertrial reinforcement. This could indicate that a few of the N_1, N_2, N_3, $\cdots$, are not removed entirely by the intertrial reinforcement, and after a sufficiently large number of trials, enough of them have become conditioned to running to increase the resistance to extinction. Another possible interpretation is that after a large number of trials, a new set of elements, such as frustration (Amsel, 1958, 1962, 1967), may come into play. These elements would not be present on every nonreinforced trial (those early in acquisition) but only appear, to be conditioned and to increase resistance to extinction, following a large number of trials.

Twenty-four-hour intertrial interval. Capaldi and Capaldi (1970), Capaldi and Spivey (1964a, 1964b), and Weinstock (1954, 1958) have found the partial reinforcement effect with a 24-hour intertrial interval. The only reason this is not accounted for by the model is that, of course, the subjects are fed each day, so it becomes the same problem as intertrial reinforcement above and can be handled in the same way. However, since the Capaldi and Capaldi (1970) experiment used a relatively small number of trials, it may be that the present model, which is a modification of Estes (1959) stimulus fluctuation model, would call on some of the properties of the original model. That is, over the 24-hour period the stimulus is modified by random fluctuation so that in the partially reinforced group some of the N_1, N_2, N_3, $\cdots$, elements come back into the sample.

Effort. Lawrence and Festinger (1962) found that increasing the effort required to make the response in partially reinforced groups increased resistance to extinction. Those subjects in a situation requiring much effort to respond have a larger population of stimuli to sample because the increased effort produces more (proprioceptive) stimulus elements to which running can be conditioned. The low effort groups do not have these additional stimuli present during learning and, thus, they have running conditioned to fewer elements and show less resistance to extinction.

LIMITATIONS OF THE MODEL

Although the model integrates a large amount of data, there are two points at which the model does not fit the data well with some parameters. Both of these are shown in Figure 5.

Low probability of response for partially reinforced groups during acquisition. Although the groups given partial reinforcement have more elements conditioned than the continuously reinforced group, the proportion of these elements conditioned to responding on most acquisition trials is smaller in the partially reinforced schedules. Most partially reinforced groups learn nearly as rapidly as continuously reinforced groups, at least their speed of response is not as low as would be implied by the model predictions represented in Figure 5.

Sudden drop in probability of response for the continuously reinforced group during extinction. If the N_1, N_2, N_3, $\cdots$, are very large in relation to N_0, the continuously reinforced group which has been responding with probability near 1.0 will show too rapid a decrease in probability of response during extinction as shown in Figure 5. For instance, if $N_0 = N_1$, probability of response on the first extinction trial will be near 1.0, but probability of response will drop to less than .5 on the second extinction trial. The larger N_0 is relative to N_T, the slower will be this drop in probability of response during the early trials of extinction. The parameters used here were $N_0 = 200$,

$N_T = 1{,}000$, so that $N_1 = 160$; thus with an N_0 this small relative to N_T, the drop is quite rapid as shown in Figure 5.

These limitations could probably be corrected by adding another sampling assumption making sampling an active process with an element more likely to be sampled the more times it was reinforced, rather than having all stimuli available being sampled on each trial. This would be similar to the stimulus analyzers proposed by Sutherland et al. (1965) except that rather than talking about analyzers being tuned in, in a stimulus-sampling model one would talk about elements becoming more likely to be sampled. In the first case, the partially reinforced groups would begin sampling from a relatively limited number of the stimuli present after being on the schedules for some time so their probability of response would not be so low. In the second case, the 100% group would be likely to sample from the N_0 elements for several trials during extinction so probability of response would not drop as suddenly when a large number of N_1, N_2, N_3, . . ., elements are suddenly included in the sample.

This assumption has not been included in the model as presented for two reasons: (*a*) The severity of these limitations is related to the particular parameters used, so that with some parameters these limitations nearly disappear, and (*b*) a model that adds a new assumption to account for each new finding soon becomes only a listing of relevant variables rather than an integrating and predicting model.

Of course, this model is closely related to Capaldi's (1966) hypothesis of sequential effects in that it is essentially a statement of that theory in stimulus-sampling terms. This model also deals with the effect of stimulus modification on resistance to extinction and emphasizes the sequential variables of N-length, number of different N-lengths, and number of occurrences of each N-length. By stating the theory in stimulus-sampling terms and constructing a computer simulation model, the theory is shown to account for such things as the pattern of response speeds in single-alternation acquisition. Also, stated as a stimulus-sampling model, it can be more easily manipulated by adding assumptions to account for data not specifically accounted for in the model as it stands. Finally, the model now makes unambiguous predictions that were not possible with it stated only verbally, such as the question of whether 10 runs of N-length 5 would result in more or less resistance to extinction than 15 runs of N-length 3 in a given situation.

The model can also accommodate findings that Amsel's (1958, 1962, 1967) frustration theory can, if anticipatory frustration is viewed as a set of stimulus elements. For instance, the finding that there is increased resistance to extinction following a large number of trials but not after a small number of trials when intertrial reinforcement is given is consonant with the model if anticipatory frustration is seen as consisting of stimulus elements which are not present early in acquisition, but appear later in acquisition. Thus, when few trials are run with intertrial reinforcement, running is not conditioned to these frustration elements; if many trials are given, these frustration elements are present during acquisition and running becomes conditioned to them.

Of course, the model is basically a modification of the N-element pattern model of Atkinson, Bower, and Crothers (1965), and of the stimulus-modification model of Estes' (1959) in which the sources of stimuli in the situation are not only undergoing random fluctuation, but some of them are under the control of the reinforcement schedule. That is, rather than a constant random fluctuation taking place, after one nonreinforced trial certain stimulus elements (N_1) are brought into the sample; after two nonreinforced trials in succession more elements (N_2) are brought into the sample, etc.

This model now makes it possible for stimulus-sampling theory to account for the increased resistance to extinction following partial reinforcement, a problem with which Estes (1959) struggled, but for which he never arrived at an adequate solution. It not only makes stimulus-sampling theory more complete but makes it possible to predict resistance to extinction in a group given any schedule of reinforcement.

REFERENCES

AMSEL, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, 1958, **55**, 102–119.

AMSEL, A. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological Review*, 1962, **69**, 306–328.

AMSEL, A. Partial reinforcement effects on vigor and persistence: Advances in frustration theory derived from a variety of within-subjects experiments. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1967.

AMSEL, A., HUG, J. J., & SURRIDGE, C. T. Number of food pellets, goal approaches, and the partial reinforcement effect after minimal acquisition. *Journal of Experimental Psychology*, 1968, **77**, 530–534.

ATKINSON, R. C., BOWER, G. H., & CROTHERS, E. J. *An introduction to mathematical learning theory*. New York: Wiley, 1965.

BACON, W. E. Partial-reinforcement extinction following different amounts of training. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 998–1003.

BERG, R. F., CONNELLY, J. F., & DUERFELDT, P. H. Replacement function of intertrial reinforcement. *Psychonomic Science*, 1969, **15**, 37–38.

BLACK, R. W., & SPENCE, K. W. Effects of intertrial reinforcement on resistance to extinction following extended training. *Journal of Experimental Psychology*, 1965, **70**, 559–563.

BLOOM, J. M. Early acquisition responding on trials following different rewards and nonrewards. *Psychonomic Science*, 1967, **7**, 37–38.

BLOOM, J. M., & CAPALDI, E. J. The behavior of rats in relation to complex patterns of partial reinforcement. *Journal of Comparative and Physiological Psychology*, 1961, **54**, 261–265.

BOWEN, J., & McCAIN, G. Occurrence of the partial reinforcement extinction effect after only one NRNR sequence of trials. *Psychonomic Science*, 1967, **9**, 15–16.

BROWN, R. T., & LOGAN, F. A. Generalized partial reinforcement effect. *Journal of Comparative and Physiological Psychology*, 1965, **60**, 64–69.

CAMPBELL, P. E., KNOUSE, S. B., & WROTEN, J. D. Resistance to extinction in the rat following regular and irregular schedules of partial reward. *Journal of Comparative and Physiological Psychology*, 1970, **72**, 210–215.

CAPALDI, E. J. Effect of N-length, number of different N-lengths, and number of reinforcements on resistance to extinction. *Journal of Experimental Psychology*, 1964, **68**, 230–239.

CAPALDI, E. J. Partial reinforcement: A hypothesis of sequential effects. *Psychological Review*, 1966, **73**, 459–477.

CAPALDI, E. J., & CAPALDI, E. D. Magnitude of partial reward, irregular reward schedules, and a 24-hour ITI: A test of several hypotheses. *Journal of Comparative and Physiological Psychology*, 1970, **72**, 203–209.

CAPALDI, E. J., & HART, D. Influence of a small number of partial reinforcement training trials on resistance to extinction. *Journal of Experimental Psychology*, 1962, **64**, 166–171.

CAPALDI, E. J., HART, D., & STANLEY, L. R. Effect of intertrial reinforcement on the aftereffect of nonreinforcement and resistance to extinction. *Journal of Experimental Psychology*, 1963, **65**, 70–74.

CAPALDI, E. J., & KASSOVER, K. Sequence, number of nonrewards, anticipation, and intertrial interval in extinction. *Journal of Experimental Psychology*, 1970, **84**, 470–476.

CAPALDI, E. J., & MINKOFF, R. Influence of order of occurrence of nonreward and large and small reward on acquisition and extinction. *Journal of Experimental Psychology*, 1969, **81**, 156–160.

CAPALDI, E. J., & SPIVEY, J. E. Effect of goalbox similarity on the aftereffect of nonreinforcement and resistance to extinction. *Journal of Experimental Psychology*, 1963, **66**, 461–465.

CAPALDI, E. J., & SPIVEY, J. E. Intertrial reinforcement and aftereffects at 24-hour intervals. *Psychonomic Science*, 1964, **1**, 181–182. (a)

CAPALDI, E. J., & SPIVEY, J. E. Stimulus consequences of reinforcement and nonreinforcement: Stimulus traces or memory. *Psychonomic Science*, 1964, **1**, 403–404. (b)

CAPALDI, E. J., & WARGO, P. Effect of transitions from nonreinforced to reinforced trials under spaced-trial conditions. *Journal of Experimental Psychology*, 1963, **65**, 318–319.

CAPALDI, E. J., & Waters, R. W. Conditioning and nonconditioning interpretations of small-trial phenomena. *Journal of Experimental Psychology*, 1970, **84**, 518–522.

COUGHLIN, R. C. Frustration effect and resistance to extinction as a function of percentage of reinforcement. *Journal of Experimental Psychology*, 1970, **84**, 113–119.

ERLEBACHER, A. Reversal learning in rats as a function of percentage of reinforcement and degree of learning. *Journal of Experimental Psychology*, 1963, **66**, 84–90.

ESTES, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. *General systematic formulations, learning, and special processes*. New York: McGraw-Hill, 1959.

GALBRAITH, K., RASHOTTE, M. E., & AMSEL, A. Within-subjects partial reinforcement effects varying percentage of reward to the partial stimulus between groups. *Journal of Experimental Psychology*, 1968, **77**, 547–551.

GONZALEZ, R. C., & BITTERMAN, M. E. Resistance to extinction in the rat as a function of percentage and distribution of reinforcement. *Journal of Comparative and Physiological Psychology*, 1964, **58**, 258–263.

GROSSLIGHT, J. H., & RADLOW, R. Patterning effect of the nonreinforcement–reinforcement sequence in a discrimination situation. *Journal of Comparative and Physiological Psychology*, 1956, **49**, 542–546.

GROSSLIGHT, J. H., & RADLOW, R. Patterning effect of the nonreinforcement–reinforcement sequence involving a single nonreinforced trial. *Journal of Comparative and Physiological Psychology*, 1957, **50**, 23–25.

HABU, Y., & ONO, S. An information-theoretic approach to partial reinforcement effect: Determination of the effective order of entropy associated with reinforcement pattern. *Japanese Psychological Research*, 1969, 11, 76–87.

HILL, W. F., & SPEAR, N. E. Extinction in a runway as a function of acquisition level and reinforcement percentage. *Journal of Experimental Psychology*, 1963, **65**, 495–500.

JERNSTEDT, G. C. Joint effects of pattern of reinforcement, intertrial interval, and amount of reinforcement in the rat. *Journal of Comparative and Physiological Psychology*, 1971, **75**, 421–429.

KOTESKEY, R. L., & HENDRIX, M. M. Increased resistance to extinction as a function of double and single alternation and of subsequent continuous reinforcement. *Journal of Experimental Psychology*, 1971, **88**, 423–428.

KOTESKEY, R. L., & STETTNER, L. J. The role of nonreinforcement–reinforcement sequences in the partial-reinforcement effect. *Journal of Experimental Psychology*, 1968, **76**, 198–205.

LAWRENCE, D. H., & FESTINGER, L. *Deterrents and reinforcement.* Stanford: Stanford University Press, 1962.

LEWIS, D. J. Partial reinforcement: A selective review of the literature since 1950. *Psychological Bulletin*, 1960, **57**, 1–28.

MCCAIN, G. Partial reinforcement with a small number of trials: Performance in extinction. *Psychonomic Science*, 1965, **2**, 131–132. (a)

MCCAIN, G. Partial reinforcement with a small number of trials: Modified extinction procedures. *Psychonomic Science*, 1965, **2**, 133–134. (b)

MCCAIN, G. Partial reinforcement with a small number of acquisition trials: Pretraining and acquisition. *Psychonomic Science*, 1965, **2**, 249–250. (c)

MCCAIN, G. Partial reinforcement effects following a small number of acquisition trials. *Psychonomic Monograph Supplements*, 1966, 1(Whole No. 12).

MCCAIN, G. The partial reinforcement effect after minimal acquisition: Single pellet reward. *Psychonomic Science*, 1968, **13**, 151–152.

MCCAIN, G., & BROWN, E. R. Partial reinforcement with a small number of trials: Two acquisition trials. *Psychonomic Science*, 1967, **7**, 265–266.

PAVLIK, W. B., & CARLTON, P. L. A reversed partial-reinforcement effect. *Journal of Experimental Psychology*, 1965, **70**, 417–423.

RASHOTTE, M. E., & AMSEL, A. The generalized PRE: Within-S PRF and CRF training in different runways, at different times of day, by different experimenters. *Psychonomic Science*, 1968, **11**, 315–316.

RASHOTTE, M. E., ROSS, M., & AMSEL, A. Generalization of the partial reinforcement effect. *Psychonomic Science*, 1968, **11**, 173–174.

RUDY, J. W., HOMZIE, M. J., COX, R. D., GRAEBER, R. C., & CARTER, E. N. Effects of sequential manipulations in the within-subjects partial reinforcement runway experiment with rats. *Journal of Comparative and Physiological Psychology*, 1970, **72**, 105–115.

SPENCE, K. W., PLATT, J. R., & MATSUMOTO, R. Intertrial reinforcement and the partial reinforcement effect as a function of number of training trials. *Psychonomic Science*, 1965, **3**, 205–206.

SPIVEY, J. E. Resistance to extinction as a function of N-R transitions and percentage of reinforcement. *Journal of Experimental Psychology*, 1967, **75**, 43–48.

SPIVEY, J. E., & HESS, D. T. Effect of partial reinforcement trial sequences on extinction performance. *Psychonomic Science*, 1968, **10**, 375–376.

SURRIDGE, C. T., RASHOTTE, M. E., & AMSEL, A. Resistance to extinction of a running response after a small number of partially rewarded trials. *Psychonomic Science*, 1967, **7**, 31–32.

SUTHERLAND, N. S., MACKINTOSH, N. J., & WOLFE, J. B. Extinction as a function of the order of partial and consistent reinforcement. *Journal of Experimental Psychology*, 1965, **69**, 56–59.

THEIOS, J. The partial reinforcement effect sustained through blocks of continuous reinforcement. *Journal of Experimental Psychology*, 1962, **64**, 1–6.

THEIOS, J., & MCGINNIS, R. W. Partial reinforcement before and after continuous reinforcement. *Journal of Experimental Psychology*, 1967, **73**, 479–481.

TYLER, D. W., WORTZ, E. C., & BITTERMAN, M. E. The effect of random and alternating partial reinforcement on resistance to extinction in the rat. *American Journal of Psychology*, 1953, **66**, 57–65.

WEINSTOCK, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *Journal of Comparative and Physiological Psychology*, 1954, **47**, 318–322.

WEINSTOCK, S. Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *Journal of Experimental Psychology*, 1958, **56**, 151–158.

WILSON, J. J. Level of training and goal-box movements as parameters of the partial reinforcement effect. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 211–213.

(Received November 1, 1971)

Psychological Review
1972, Vol. 79, No. 2, 172–177

THEORETICAL NOTES

CAN A TWO-STATE MODEL ACCOUNT FOR TWO-CHOICE REACTION-TIME DATA?[1]

JOHN THEIOS[2] AND PETER G. SMITH

University of Wisconsin

A class of two-state models involving the notion of a stimulus bias is outlined for the stimulus-identification–response-selection operation in two-choice reaction-time (2CRT) tasks. The model is applied to sequential RT data from a 2CRT situation, and the agreement between predictions and the data is reasonable. From consideration of other data, a case is made that sequential effects in 2CRT are due to the structure of the sequence of stimuli, rather than due to local response biasing resulting from the sequence of required responses.

A number of investigators (e.g., Keele, 1969; Laming, 1968; Remington, 1969; Schvaneveldt & Chase, 1969) have found reliable sequential effects in two-choice reaction-time (2CRT) tasks using relatively long interstimulus intervals and a compatible stimulus–response arrangement. The most extensive analysis of 2CRT sequential effects has been made by Remington (1969). He found systematic and orderly stimulus sequential dependencies up to the fifth order in mean reaction times (RTs). On the basis of these high-order sequential effects, he concludes that simple two-state models, such as that proposed by Falmagne (1965), are inadequate as representations of the 2CRT process. In his own words, Remington states:

> It appears that premature theorizing based upon the repetition effect has been relatively ineffective in producing models with adequate explanatory and predictive power. . . . Falmagne's mathematical model is capable of making relatively good predictions concerning many interesting aspects of CRT data. However, there is convincing evidence that the processes underlying choice reaction behavior are too complex to be accurately represented by such a two-state model [pp. 256–257].

While Remington's conclusion is true for some situations (cf. Falmagne, 1968) and we generally agree with it (Falmagne & Theios, 1969), the conclusion certainly cannot be drawn from 2CRT data. The purpose of this note is to show that the contrary seems to be the case; 2CRTs can be accurately represented as a two-state process.

CLASS OF TWO-STATE MODELS

Preparation Model

Falmagne (1965) assumed that at any moment in time, the subject is either prepared or unprepared to make the response associated with each stimulus in a CRT experiment and that trial-by-trial changes in the state of preparation are governed by the following four assumptions: If the subject is prepared to make response r_i to stimulus s_i on Trial n, and if stimulus s_i is presented, then RT will be short (a random sample from a distribution with a relatively small mean) and on the next trial the subject will still be prepared to make r_i to stimulus s_i with probability 1. If the subject is prepared for $s_i - r_i$ on Trial n and if stimulus s_i is *not* presented, then on the next trial the subject will still be prepared for $s_i - r_i$ with probability $1 - c'$ and will be unprepared for $s_i - r_i$ with probability c'. If the subject is not prepared to make response r_i to stimulus s_i on Trial n, and if stimulus s_i is presented, then RT will be long (a random sample from a distribution having a rather large mean), and on the next trial, the subject will be prepared for $s_i - r_i$ with probability c and not prepared for $s_i - r_i$ with probability $1 - c$. If the subject is not prepared for $s_i - r_i$ on Trial n,

---

[1] This research was supported by Public Health Service Research Grants MH 16711-01 and MH 19006-01 from the National Institute of Mental Health, and by grants from the Wisconsin Alumni Research Foundation. Use of the University of Wisconsin Computing Center was made possible through support, in part, from the National Science Foundation. This report was prepared during the senior author's tenure as a John Simon Guggenheim Memorial Fellow visiting at the Laboratory of Experimental Psychology, the Free University of Brussels, Brussels, Belgium. We wish to thank Paul Bertelson, the director of the Laboratory, for his generous hospitality.

[2] Requests for reprints should be sent to John Theios, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

and if stimulus s_i is not presented, then on the next trial the subject will still be unprepared for $s_i - r_i$ with probability 1.

For an experiment in which stimulus s_i is presented with a constant probability π_i, the four assumptions imply that the trial-to-trial states of preparation to $s_i - r_i$ form a two-state ergodic Markov chain in which the asymptotic probability (P_i) of being prepared for $s_i - r_i$ is

$$P_i = (\pi_i c)/[\pi_i c + (1-\pi_i)c']. \qquad [1]$$

The assumption that the states of preparation for all the stimuli in the experimental set are independent of eacho ther implies the counterintuitive state of affairs that on any trial the subject may be prepared for all, some, or none of the stimuli! However, if the assumption is made that $c' = c$, then the model requires the subject to be prepared for only one $s_i - r_i$ pair on any trial, and it predicts that the asymptotic probability that the subject is prepared for $s_i - r_i$ is simply π_i (the probability matching prediction).

Memory Stack Model

When applied to a 2CRT experiment, the memory search model suggested by Falmagne and Theios (1969) reduces to a simple hierarchy or stack of two $s - r$ associations which is searched in a serial, self-terminating manner. For two stimuli (s_1 and s_2), there are two states of bias, B_1 and B_2: B_1 represents the state of affairs when s_1 is in the first memory position and s_2 is in the second memory position; B_2 represents the state of affairs when s_1 is in the second memory position and s_2 is in the first memory position. When stimulus s_i ($i=1$ or 2) is presented, there is a probability c_i that the stored representation of $s_i - r_i$ takes the first memory position if it was not already stored there. This assumption leads to the simple Markov process depicted in Figure 1.

Following Falmagne and Theios (1969), it may be assumed that with each of the two states of the process, there is associated a normal distribution of stimulus identification—RTs such that the mean and variance for the first state are less than those associated with the second state. If s_1 (or s_2) is presented and B_1 (or B_2) is the state of the process, then the RT will be a random sample from a normal distribution with parameters μ_1 and σ_1, and the RT will generally be relatively short. If s_1 (or s_2) is presented and B_2 (or B_1) is the state of the process, then the RT will be a random sample from a normal distribution with parameters μ_2 and σ_2, and the RT will be relatively long. The difference $\mu_2 - \mu_1$ may be interpreted as the average time to make a comparison between

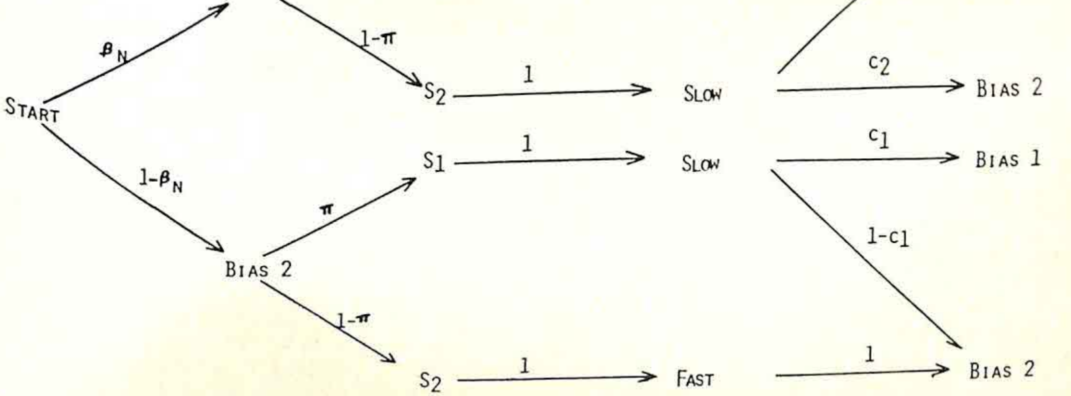


FIG. 1. Two-state ergodic Markov process for 2CRTs. (Short [or long] RTs may be considered as random samples from a normal distribution having a mean μ_1 [or μ_2].)

the representation of the presented stimulus and the contents of the second memory position. The normality assumptions are tantamount to assuming that, trial by trial, the additive components of stimulus input time, memory search and comparison time, and response execution time are each determined by multiple factor subprocesses that, when convoluted, together yield a normal distribution for a large sample of trials. The assumption of normal distributions, although traditional, is not crucial to the treatment presented here, and any reasonable distributions could be substituted without significantly changing the theory.

Assuming only that there are means μ_1 and μ_2 associated with the first and second states, respectively, t_n, the expected RT on any Trial n, will be a simple binary probability mixture (Falmagne, 1965, 1968)

$$E(t_n) = (\beta_n)\mu_1 + (1 - \beta_n)\mu_2. \qquad [2]$$

In Equation 2, β_n represents the conditional probability of the representation of the stimulus presented on Trial n being in the first memory position given the previous sequence of stimulus events on all trials prior to n. The model accounts for sequential dependencies in RT by averaging short and long RTs in different proportions, depending on the exact sequence of preceding stimuli. As n becomes large, the model predicts that the asymptotic value of β will be

$$\beta = (\pi c_1)/[\pi c_1 + (1 - \pi)c_2]. \qquad [3]$$

If it is assumed that $c_1 = c_2 = c$, then the memory stack model makes the same predictions as Falmagne's (1965) preparation model with $c' = c$, including the asymptotic probability matching prediction that $\beta = \pi$.

ANALYSIS OF TWO-CHOICE RT SEQUENTIAL DEPENDENCIES

Remington (1969) used stimulus presentation probabilities (π, $1 - \pi$) of (.5, .5) and (.3, .7) in a 2CRT experiment. The stimuli consisted of left and right lights (e.g., s_1 and s_2) and the corresponding responses were a left- and a right-hand index finger button push (e.g., r_1 and r_2). Each trial began with a red warning light followed by a one-second foreperiod before the presentation of s_1 or s_2. The interstimulus interval averaged about four seconds. Each of five human subjects were given 1,200 trials on each problem in alternating blocks of 200 trials, and the data represent the last 800 trials for the (.5, .5) problem and the last 1,000 trials for the (.3, .7) problem.

Application of the Two-State Model

We applied the two-state model represented in Figure 1 to the Remington (1969) data as presented in his Figures 2 and 3. There are four parameters to the model: c_1 and c_2 are the respective probabilities of representations of stimuli s_1 and s_2 entering the first memory position after they are presented, and μ_1 and μ_2 are the means of the RT distributions associated with finding a matching stimulus representation in the first and second memory positions, respectively. In Figure 1, β_n represents the probability of stimulus s_1 being in the first memory position on Trial n, and it will be taken as its asymptotic value predicted by Equation 3. Two theoretical hypotheses were investigated:

Hypothesis 1 (probability matching) requires that the value of c for all three stimulus probability conditions (.3, .5, .7) be the same. Hypothesis 1 assumes that, given it is presented, each stimulus has an equal probability of entering the first state independent of its presentation probability and that the mean RTs for the first and second states are independent of the presentation probabilities of the stimuli whose representations are currently in the states. Three parameters were estimated under Hypothesis 1: c, μ_1, and μ_2.

Hypothesis 2 permits each of the conditions to have its own individual c value. Hypothesis 2 assumes that, given it is presented, the probability that a stimulus representation enters the bias state is a function of its presentation probability, but that the fastness or slowness of the response depends only on the position of the matching representation and does not depend on the presentation probability of the stimulus. Five parameters were estimated under Hypothesis 2: $c(.3)$, $c(.5)$, $c(.7)$, μ_1, and μ_2.

Chandler's (1969) subroutine STEPIT was used to find values of the parameters under the two hypotheses which would simultaneously minimize the weighted root mean squared error (cf. Sternberg, 1969) between Remington's (1969) 93 conditional mean RTs[3] and the corre-

---

[3] In Remington's (1969) data reproduced in Table 2, the sequential fifth-order conditional mean RTs are independent observations. However, all lower order conditional mean RTs may be obtained as weighted averages of higher order conditional mean RTs. Thus, to the extent that the higher order data are in agreement with the theoretical predictions, the lower order data will also be in agreement.

TABLE 1

TWO-STATE MODEL PARAMETERS FOR REMINGTON (1969) RT DATA

Parameter	π	Hypothesis 1 (c's equal)	Hypothesis 2 (c's free)
c	.3	.51	.18
	.5	.51[a]	.22
	.7	.51[a]	.38
μ_1		266	263
μ_2		311	317
Root mean squared error		8.6 msec.	2.8 msec.

Note.—π = stimulus presentation probability; c = state transition probability; μ_1 = mean RT in bias state; μ_2 = mean RT in nonbias state.

[a] Parameter value not estimated, but fixed by theoretical considerations.

sponding conditional expected RTs predicted by the two-state model depicted in Figure 1. The minimum root mean squared errors and the corresponding parameter estimates are given in Table 1. Hypothesis 1 yielded an average error of prediction of 8.6 milliseconds (msec.). Hypothesis 2 yielded an average absolute error of less than a third that of Hypothesis 1. The observed and predicted (Hypothesis 2) conditional mean reaction times are presented in Table 2. The predictions follow the observed sequential dependencies very well. The average absolute error of prediction is only 2.8 msec. In summary, the two-state model does a remarkable job of accounting for sequential effects in the 2CRT situation.

FURTHER EVIDENCE FOR A TWO-STATE PROCESS

A common practice in some CRT experiments is to have the subject predict what stimulus he thinks will occur on the next trial. In these prediction experiments it has been found consistently that RT is shorter when the stimulus is correctly anticipated than when the stimulus is not expected (Bernstein & Reese, 1965, 1967; Bernstein, Schurman, & Forester, 1967; Keele, 1969). For example, Hinrichs (1970) ran a 2CRT prediction experiment with stimulus probabilities of $\frac{2}{3}$ and $\frac{1}{3}$. He found: (*a*) subjects' predictions matched the stimulus probabilities, (*b*) RTs following *correct* predictions were virtually identical for the more frequent (424 msec.) and the less frequent (422 msec.) stimuli, (*c*) RTs following *incorrect* predictions did not differ significantly between the more frequent (474 msec.) and less frequent (486 msec.) stimuli, (*d*) the mean RT to the more frequent stimulus did not differ significantly for the subjects who were required to make a prediction (446 msec.) and subjects who were not required to make predictions (440 msec.), (*e*) the mean RT to the less frequent stimulus did not differ significantly for the predict group (463 msec.) and the nonpredict group (473 msec.). These results can be interpreted in terms of the present theoretical framework by assuming that the subject's prediction corresponds to the contents of the first memory position. If the subject correctly predicts the stimulus, his RT to it is short because its representation was in the first position and the identification match was made quickly. If the subject incorrectly predicts the stimulus, its representation was not in the first position, and his RT was long because two comparison operations had to be carried out before a correct identification match could be obtained.

Two alternative elaborations of this interpretation can be considered. The first is that of a passive system in which the prediction response is simply a read-out of the contents of the first memory position which is determined stochastically by the sequence of stimuli. The alternative assumption is that the act of predicting is under cognitive control and, in the interest of consistency, the first memory position is loaded with the stimulus representation corresponding to the subject's prediction.

STIMULUS BIAS VERSUS RESPONSE BIAS

The Falmagne (1965) and Falmagne and Theios (1969) theories clearly require stimulus biases on the part of the subject, with no theoretical provision being made for response biases. This presents no problems when there is a one-to-one mapping from the set of stimuli to the set of responses (as in 2CRT experiments), since a stimulus bias cannot be separated from a response bias with a one-to-one mapping. However, the Remington (1969) sequential effects and the results of the many prediction-RT experiments cited above could be due to a response bias process rather than a stimulus bias process. To determine whether the prediction-RT results were due to a stimulus as opposed to response bias, Hinrichs and Krainz (1970) adapted the LaBerge-Tweedy (1964) three-stimulus–two-response RT design to the prediction situation. Stimulus *A* was assigned to one response (R_A) and Stimuli B_1 and B_2 were assigned to the other response

TABLE 2

MEAN REACTION TIME (IN MILLISECONDS) ON TRIAL n GIVEN PRECEDING STIMULUS SEQUENCE

Order	Stimulus sequence					Stimulus presentation probability					
						.3		.7		.5	
	s_n	s_{n-1}	s_{n-2}	s_{n-3}	s_{n-4}	Obs.	Pre.	Obs.	Pre.	Obs.	Pre.
First	1					307	308	274	273	290	290
Second	1	1				302	300	269	269	284	284
	1	2				310	311	280	281	295	296
Third	1	1	1			297	293	267	267	278	280
	1	1	2			309	304	276	274	288	289
	1	2	1			306	306	278	278	290	292
	1	2	2			314	314	292	288	300	301
Fourth	1	1	1	1		292	288	263	266	276	276
	1	1	1	2		301	295	271	270	280	284
	1	1	2	1		309	298	275	272	283	286
	1	1	2	2		309	304	279	278	293	293
	1	2	1	1		309	302	278	276	292	288
	1	2	1	2		305	308	280	282	286	295
	1	2	2	1		308	309	291	285	293	296
	1	2	2	2		316	315	294	293	305	304
Fifth	1	1	1	1	1	287	283	264	265	273	274
	1	1	1	1	2	296	290	262	268	279	279
	1	1	1	2	1	296	292	273	269	277	281
	1	1	1	2	2	302	297	266	273	283	286
	1	1	2	1	1	303	295	273	271	285	283
	1	1	2	1	2	311	300	278	275	281	288
	1	1	2	2	1	309	302	275	277	293	290
	1	1	2	2	2	309	305	285	282	294	295
	1	2	1	1	1	303	300	279	275	292	285
	1	2	1	1	2	312	304	276	279	292	291
	1	2	1	2	1	314	306	283	281	282	292
	1	2	1	2	2	300	309	276	286	290	298
	1	2	2	1	1	308	308	287	284	288	294
	1	2	2	1	2	308	312	300	289	300	300
	1	2	2	2	1	320	313	288	291	301	301
	1	2	2	2	2	315	316	302	297	310	307

Note.—Obs. = observed; Pre. = predicted.

(R_B). Each of the three stimuli (which were visual displays of the digits 0, 1, and 2) occurred in a random sequence with probability $\frac{1}{3}$. Twenty-four university students were given 140 trials in which they were required to predict a stimulus before one was presented for their reaction. The data from the last 90 trials of the Hinrichs and Krainz (1970) experiment were analyzed. The results are consistent with the earlier two-stimulus prediction-RT experiments: (*a*) If the presented stimulus was correctly predicted, the RT was short, having a mean of 485 msec.; (*b*) if the presented stimulus was not anticipated by the subject, the RT was long, having a mean of about 565 msec. Further, the important findings relating to the question of stimulus versus response bias were: (*c*) the RTs to *expected* stimuli were the same, 485 msec., independent of whether the response required was R_A which occurred on $\frac{1}{3}$ of the trials or R_B which occurred on $\frac{2}{3}$ of the trials; (*d*) there was no significant difference between the RT to an *unexpected* stimulus as a function of whether a $\frac{1}{3}$ probability R_A response (548 msec.) or a $\frac{2}{3}$ probability R_B response (567 msec.) was required. In summary, there is absolutely no evidence in the Hinrichs and Krainz (1970) data to support a response bias hypothesis.

Most importantly, the data from the Hinrichs and Krainz (1970) experiment are entirely consistent with the two-state model presented in Figure 1. The prediction made by the subject indicates the contents of the first memory position, and the conditional RT is simply short or long, depending on whether or not the contents of the first memory position matched the presented stimulus. This result is independent of the 2:1 response requirement favoring *B* responses. In addition, the subjects' predictions closely matched the stimulus probabilities as required by the two-state model.

REFERENCES

BERNSTEIN, I. H., & REESE, C. Behavioral hypotheses and choice reaction time. *Psychonomic Science*, 1965, **3**, 259–260.

BERNSTEIN, I. H., & REESE, C. Choice reaction time and behavioral hypotheses: The effects of learning sequential redundancies. *Psychonomic Science*, 1967, **9**, 189–190.

BERNSTEIN, I. H., SCHURMAN, D. L., & FORESTER, G. Choice reaction time as a function of stimulus uncertainty, response uncertainty, and behavioral hypotheses. *Journal of Experimental Psychology*, 1967, **74**, 517–524.

CHANDLER, J. P. STEPIT: Finds local minima of a smooth function of several parameters. (CPA 312) *Behavioral Science*, 1969, **14**, 81–82.

FALMAGNE, J. C. Stochastic models for choice reaction time with applications to experimental results. *Journal of Mathematical Psychology*, 1965, **12**, 77–124.

FALMAGNE, J. C. Note on a simple fixed-point property of binary mixtures. *British Journal of Statistical and Mathematical Psychology*, 1968, **21**, 131–132.

FALMAGNE, J. C., & THEIOS, J. On attention and memory in reaction time experiments. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North Holland Publishing, 1969.

HINRICHS, J. V. Probability and expectancy in two-choice reaction time. *Psychonomic Science*, 1970, **21**, 227–228.

HINRICHS, J. V., & KRAINZ, P. L. Expectancy in choice reaction time: Anticipation of stimulus or response? *Journal of Experimental Psychology*, 1970, **85**, 330–334.

KEELE, S. W. Repetition effect: A memory dependent process. *Journal of Experimental Psychology*, 1969, **80**, 243–248.

LABERGE, D., & TWEEDY, J. R. Presentation probability and choice time. *Journal of Experimental Psychology*, 1964, **68**, 477–481.

LAMING, D. R. J. *Information theory of choice-reaction times*. London: Academic Press, 1968.

REMINGTON, R. J. Analysis of sequential effects in choice reaction times. *Journal of Experimental Psychology*, 1969, **82**, 250–257.

SCHVANEVELDT, R. W., & CHASE, W. G. Sequential effects in choice reaction time. *Journal of Experimental Psychology*, 1969, **80**, 1–9.

STERNBERG, S. The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North Holland Publishing, 1969.

(Received May 3, 1971)

Psychological Review
1972, Vol. 79, No. 2, 178–179

CONTINUITY THEORY REVISITED:

REPLY TO D. B. BERCH[1]

GERALD B. BIEDERMAN[2]

University of Toronto

D. B. Berch has argued that evidence by G. B. Biederman that S− is nonmonotonic in habit strength should be reinterpreted on the basis of the generalization of S+ responses to S−. Berch's argument is found to be inconsistent with existing experimental evidence and fails to refute Biederman's contention that the inhibition associated with S− is nonmonotonic with respect to nonreinforced trials.

Berch (1971) has found fault with Biederman's (1970) suggestion that Spence's (1936) theory suffers a failure in the assumption of monotonicity in the habit strength of the negative stimulus (S−). Berch's main argument concerns the interpretation of experimental evidence indicating that S− is nonmonotonic in habit strength.

In the procedure first employed by Deutsch and Biederman (1965), subjects are trained in two simultaneous discrimination problems, with one of the problems receiving more training than the other. At various stages, a choice is offered between the negative stimuli from the two problems (Biederman, 1967, 1968; Deutsch & Biederman, 1965). Subjects show a significant preference for the S− from the more trained discrimination problem. This has been interpreted as indicating that S− (from the more trained problem) is less aversive than S− from the less trained problem, suggesting that the aversiveness of S− is not monotonic. Berch raises an interpretation that subjects are exhibiting generalized approach to the more trained problem. This interpretation is not responsive to the fact that when probe trials are presented later in training, choice becomes random between more and less trained S− (Biederman, 1967, 1968). In Biederman (1968), subjects strongly preferred the more trained S− after 150 total trials (100 more trained, 50 less trained), but probes at 300 and 450 total trials were not significant. If subjects were choosing S− on the basis of generalization of approach to S+ (i.e., showing a preference for the more trained problem per se), choices at later stages of training must similarly favor the more trained S−, inasmuch as performance in the more trained discrimination task is superior to performance in the less trained problem later in training (Biederman, 1967, 1968). If S− is nonmonotonic, the preference for the more trained S− should fade as the less trained S− similarly loses its aversiveness.

Berch's generalization interpretation of the early preference for the more trained discrimination problem may not be consistent with existing data. In absolute terms, subjects make more errors in the more trained situation (Biederman, 1967; Deutsch & Biederman, 1965). In Biederman (1968), for example, for the first 150 total trials, subjects made an average of 30.5 errors (69.5% correct) in the 100 more trained trials, and 19.75 errors (60.5% correct) in the 50 less trained trials.

It is not clear that invoking a generalization process forces the prediction that the more trained is more attractive than the less trained problem; for this to be the case (assuming S− is monotonic), S− must have some degree of inhibition which does not outweigh the "generalized habit (approach) tendency" of the more trained problem. Considering the absolute amount of nonreinforcement present in the three studies mentioned for more trained S−, the prediction that the more trained discrimination would still show generalized approach to a greater degree than the less trained task seems problematic.

In the absence of the probe trial data which show a highly significant choice of more trained S− at early stages in learning, it seems improbable that choice of the more trained S− would have been predicted from an assumption of monotonicity of the inhibitory properties of S−.

There is some more direct evidence that the more trained discrimination problem is not

---

[1] Research was supported by grants from the Canadian National Research Council and from the Ontario Mental Health Foundation. I am indebted to J. A. Deutsch and W. O'Neill for reading and commenting on an earlier version of this reply.

[2] Requests for reprints should be sent to Gerald B. Biederman, Scarborough College, University of Toronto, West Hill, Ontario, Canada.

preferred to the less trained. In Biederman (1968), it was found that the *less* trained S+ is chosen in preference to the more trained S+, after 150 total trials (100 more trained, 50 less trained, $p < .008$). This suggests that the "approach" curve to S+ may be similarly nonmonotonic.

The main weakness of Berch's argument is that it is not clear what type of evidence Berch would be willing to accept which could not be dismissed out of hand by ad hoc application of the broad generalization hypothesis. The purpose of my original comments (Biederman, 1970) on Wolford and Bower (1969) was to point out that evidence for the nonmonotonicity of S− exists and theories which assume that S− is monotonic must take this evidence into consideration. Berch has failed to mention, let alone critically evaluate, independent lines of evidence supporting the nonmonotonicity position (e.g., Mandler, 1968). Berch's (1971) note does not change the status of the requirement that evidence for nonmonotonicity should be taken into consideration in constructing theories of discrimination learning. What is equally clear is that the empirical basis for the determination of the nonmonotonicity of stimulus function of S− should be broadened.

REFERENCES

BERCH, D. B. Comments on Biederman's "Continuity theory revisited: A failure in a basic assumption." *Psychological Review*, 1971, **78**, 260–261.

BIEDERMAN, G. B. The overlearning reversal effect: A function of the nonmonotonicity of the negativity of S− during discriminative training. *Psychonomic Science*, 1967, **7**, 385–386.

BIEDERMAN, G. B. Stimulus function in simultaneous discrimination. *Journal of the Experimental Analysis of Behavior*, 1968, **11**, 459–463.

BIEDERMAN, G. B. Continuity theory revisited: A failure in a basic assumption. *Psychological Review*, 1970, **77**, 255–256.

DEUTSCH, J. A., & BIEDERMAN, G. B. The monotonicity of the negative stimulus during learning. *Psychonomic Science*, 1965, **3**, 391–392.

MANDLER, J. M. Overtraining and the use of positive and negative stimuli in reversal and transfer. *Journal of Comparative and Physiological Psychology*, 1968, **66**, 110–115.

SPENCE, K. W. The nature of discrimination learning in animals. *Psychological Review*, 1936, **43**, 427–449.

WOLFORD, G., & BOWER, G. H. Continuity theory revisited: Rejected for the wrong reasons? *Psychological Review*, 1969, **76**, 515–518.

(Received June 10, 1971)

Psychological Review
1972, Vol. 79, No. 2, 180–184

COMMENTS ON "ROLES OF ACTIVATION AND INHIBITION IN SEX DIFFERENCES IN COGNITIVE ABILITIES" BY D. M. BROVERMAN, E. L. KLAIBER, Y. KOBAYSHI, AND W. VOGEL

MARY B. PARLEE[1]

Wellesley College

The hypothesis of D. M. Broverman, E. L. Klaiber, Y. Kobayshi, and W. Vogel that "known sex differences in cognitive abilities reflect sex-related differences in physiology" is criticized on the grounds that (*a*) their review and classification of the literature on sex differences in cognitive abilities is inadequate because it is selective, and arbitrarily characterizes those facts which it does include, (*b*) the pharmacological studies cited in support of their thesis involve behaviors which are at best tangentially related to human cognitive abilities and rely for their relevance on some dubious cross-species analogies, and (*c*) the authors have made, and defended, a factual error in citing evidence for an important part of their proposed mechanism by which "sex" hormones might affect autonomic activity.

Singer and Montgomery (1969) commented on one aspect of the article by Broverman, Klaiber, Kobayshi, and Vogel (1968) entitled "Roles of activation and inhibition in sex differences in cognitive abilities." In their reply to these remarks, Broverman, Klaiber, Kobayshi, and Vogel (1969) noted that Singer and Montgomery had been concerned with only part of the argument, and that they "have not challenged, or even commented upon, the major bodies of evidence presented in our paper [p. 328]." In view of the importance of and increasing work on the topic of sex differences in cognitive functioning, it seems worthwhile to consider in more detail some of the other difficulties in Broverman et al.'s (1968) paper.

Broverman et al. (1968) offer the hypothesis that "known sex differences in cognitive abilities reflect sex-related differences in physiology [p. 23]." This is a strong claim. If true, it would represent a major advance toward the goal of linking psychological functions with their physiological concomitants and would offer a partial explanation of one of the most pervasive phenomena in psychology. What follows is a brief summary of their argument, followed by a more detailed evaluation and analysis of the evidence on which it is based.

Broverman et al. (1968) begin with a review of the psychological literature on sex differences in cognitive ability and categorize the results into those tasks at which men have been found to excel and those at which women excel. They characterize the tasks on which men do better as those "more complex tasks requiring an inhibition of immediate responses to obvious stimulus attributes in favor of responses to less obvious stimulus attributes [p. 23]" and the tasks on which women's performance is better as "simple, overlearned, perceptual-motor tasks [p. 23]." The key concept of inhibition as used above is related to the "inhibitory" function of the parasympathetic division of the autonomic nervous system, and the behavioral consequences of chemically altering the balance between the sympathetic and parasympathetic nervous systems are discussed to show that parasympathetic dominance is associated with the ability to perform the inhibitory, perceptual-restructuring tasks (the kind at which men excel) and sympathetic dominance with perceptual-motor tasks (at which women surpass men). The authors then seek to establish that the "sex" hormones, androgens and estrogens, affect the balance between activation and inhibition (the vagueness of this formulation is recognized and is considered below), and they discuss possible indirect mechanisms by which androgens and estrogens might affect the autonomic nervous system (with, in their view, the androgens favoring parasympathetic and the estrogens sympathetic dominance). Thus their argument seems to be that sex differences in hormonal patterns produce sex differences in the balance of sympathetic and parasympathetic activity which produce sex differences on cognitive tasks; androgens, secreted by the male testes, favor the parasympathetic dominance necessary for the more "complex" ("inhibitory restructuring") tasks, while estrogens, secreted by the female ovaries, promote the sympathetic dominance which facilitates performance on "simple, overlearned, perceptual-motor tasks." Let us look at this argument more closely.

---

[1] Requests for reprints should be sent to Mary B. Parlee, Department of Psychology, Wellesley College, Wellesley, Massachusetts 02181.

REVIEW AND CLASSIFICATION OF LITERATURE ON SEX DIFFERENCES IN COGNITIVE ABILITIES

Of the several comprehensive reviews of the literature on human sex differences that are available (Anastasi, 1958; Garai & Scheinfeld, 1968; Maccoby, 1966; Tyler, 1965), all support the generalization that adult males perform better on mathematical and spatial tasks and adult women on verbal and perceptual-motor tasks. These reviews document the wide variety of studies on which these generalizations are based; they make no attempt, however, to abstract a factor of similarity among the types of performance at which each sex excels. Broverman et al. undertake the theoretically important task of making such an abstraction, and in so doing do not attempt to provide an exhaustive account of known sex differences, but rather a classification of certain cognitive tasks according to criteria which they have established for distinguishing between the tasks at which each sex excels. While the organization and tone of their article tend to suggest that their review of the sex differences literature is more comprehensive than it is, and that the classification of the tasks at which males and females excel was induced from the available data, it is clear from a comparison with other reviews that the studies they classify into two categories (Tables 1 and 2 of their article) were selected according to the hypotheses which they present as inductive generalizations or conclusions. In no other way could the authors arrive at a class of tasks at which females are superior which could be characterized as meeting the criteria of "simple overlearned repetitive behaviors measured in terms of speed, accuracy, or frequency of occurrence and which require minimal central [i.e., cognitive] mediation [p. 28]," since it ignores the large body of evidence (summarized in Oetzel, 1966, and elsewhere) showing that females surpass males on almost all measures of language performance. Surely it is inadequate to claim that verbal fluency, vocabulary size, grammatical sensitivity, etc., "involve minimal mediation by higher cognitive processes [p. 28]." Yet the authors make just this claim, reasoning that "Although not commonly thought of as perceptual-motor behaviors, speech and reading do involve these functions [p. 25]" and that, therefore, there is no conceptual confusion in a statement like "As with other simple perceptual-motor behaviors, females are superior to males from childhood through adulthood in verbal functions [p. 25]." Their discussion of "The behaviors [not "some of the behaviors"] in which females are superior to males . . . [p. 28]" refers to a list which includes only "reading speed."

Properly employed, the hypothetico-deductive method is essential to scientific theory construction, and its use in the Broverman et al. case is not in question. What is questionable, however, is their use of prior hypotheses (in this case the criteria for classification) to select supporting data (here, various studies of sex differences) in such a way that evidence which directly contradicts the hypotheses is ignored.

Apart from the difficulties of dealing with a restricted range of data, there is also some problem with the way Broverman et al. characterize the common factor underlying the performance of those tasks on which they report males to surpass females. They state that these tasks require "suppression of responses to immediately obvious stimulus attributes of the task in favor of responses to other, not immediately obvious, stimulus attributes . . . [p. 24]" (what these attributes—obvious or nonobvious—are is never made clear). This abstraction and description of the common underlying ability is very similar to the description by Witkin, Lewis, Hertzman, Machover, Meissner, and Wapner (1954) and Witkin, Dyk, Faterson, Goodenough, and Karp (1962) of the ability underlying males' performance on the Gottschaldt hidden figures, rod-and-frame, and other similar tasks, and Witkin et al.'s (1954) discussion of the Rorschach (inkblot) test clearly illustrates the difficulties with this characterization. Thus, while Witkin et al. (1954, p. 477) describe performance on the hidden figures test as depending on the ability to perceptually restructure ("break up") the drawing or to inhibit the response to the whole configuration (so as to concentrate on those parts relevant to the search for the hidden figure), they also speak of the subject's "passive acceptance" and response to a single feature of the inkblot rather than his "actively attempting to provide an integration of form [1954, p. 207]" of the whole figure. There seems to be an unresolved contradiction here.

Does the ability to respond to part of a complex perceptual whole involve "passive acceptance" or the presumably more active "perceptual-restructuring"? Is response to the whole configuration a lack of inhibition of the "immediately obvious" stimulus characteristics or an "integration of form"? While these questions could be put into a form which could be tested empirically, this has not yet been done, and the inconsistency which they presently illustrate brings into question the description of certain tasks as "inhibitory restructuring" without further refinement and clarification of the concept.

Thus Broverman et al.'s (1968) review and classification of the literature on sex differences in cognitive abilities seems to be selectively designed to support their hypothesis as to the criteria or attributes which distinguish cognitive tasks on which males and females excel. Their selection leads them to ignore or distort contradictory evidence (as in the case of their characterization of the tasks at which women surpass men) and to support an unclear conception of the tasks on which males are superior.

RELATION OF COGNITIVE FUNCTIONS TO THE ACTIVITY OF THE AUTONOMIC NERVOUS SYSTEM

In adducing evidence to support their proposal that "the relationship between simple perceptual-motor and inhibitory perceptual-restructuring tasks, may be, in part a function of the balance between the activating influence of the sympathetic, and the inhibitory tendencies of the parasympathetic subcortical nervous system [p. 31]," Broverman et al. (1968) rely on studies of the behavioral consequences of chemically stimulating or depressing one or the other branches of the autonomic nervous system. The logic of their argument is unexceptionable: they look at the effects of substances which stimulate/depress the sympathetic/parasympathetic nervous systems on both kinds of tasks. They seek to establish the hypotheses that "the same ergotropic functions [those mediated by the sympathetic nervous system] that facilitate wakefulness and sensory reactivity also facilitate performances of simple perceptual-motor tasks [p. 29]" and that "the same trophotropic functions [those mediated by the parasympathetic nervous system] that promote relaxation and sleep also contribute to the cognitive ability to delay initial response tendencies to obvious stimulus attributes in favor of responses to other, less obvious stimulus relationships [p. 29]." (Elsewhere these cognitive abilities which are associated with "sleep and relaxation"—as opposed to "wakefulness and sensory reactivity"—are characterized as behaviors "involving extensive mediation of higher processes" and which involve "the production of solutions to novel tasks or situations [p. 28].") The evidence for these hypotheses, however, is slim.

Even granting that some operationally clear-cut categorization of human cognitive abilities into perceptual-motor and perceptual-restructuring could be achieved, Broverman et al. in many cases fail to show that the kinds of behaviors which are affected by chemically altering the sympathetic-parasympathetic balance fall into one or the other of these categories. For example, in the section on the effects of depressing parasympathetic function on simple perceptual-motor tasks (a manipulation predicted by their hypothesis to improve performance), they cite increased open field activity, spontaneous activity in wheels, and rates of conditioned responses in rats, and motor activity in cats as "evidence . . . consistent with the proposition that an increase in the activation balance via a diminution of the inhibitory processes leads to increased output of simple perceptual-motor behaviors [p. 34]." Their notion of the appropriate extension of these sorts of animal data to human cognitive performance is illustrated elsewhere as follows: "Human females exceed human males on simple perceptual-motor tasks. The same sex difference in analogous behaviors, that is, spontaneous motor activity, is true in rats [p. 37]." Without in any way casting doubt on the relevance of animal research to human psychology, it seems fair to say that an argument that depends on an analogy between spontaneous motor activity in rats and reading speed in humans leaves much to be desired. Some examples of "perceptual-restructuring" tasks affected by chemical manipulation and cited in support of their hypothesis include alternate pushing of two levers (rats), writing slowly, temporal judgment, performance on a mirror tracing task, and "errors of overimpulsivity" in retardates.

Frequently, it seems that the studies cited by the authors in support of their hypotheses, qualify as evidence only on the assumption that the terms "activation" and "inhibition" have parallel meanings in psychology and neurophysiology; that, for example, the "inhibitory" activity of the parasympathetic nervous system is comparable to Pavlov's "conditioned inhibition," that "activation" caused by the sympathetic nervous system is reflected in spontaneous motor "activity," etc. Even granting the possibility of a simple one-to-one relation between psychological and neural processes, this sort of argument would be deficient in view of the fact that "activation" and "inhibition" can have several meanings in both psychology and neurophysiology. To take but a few examples, "activation" can refer in neurophysiology to the triggering by a neuron of other neurons or an effector, to the massive neural activity resulting from stimulation of the "reticular activating system," or to the pattern of autonomic changes resulting from stimulation of the sympathetic nervous system. In addition to the "inhibitory" action of one neuron on the firing rate of another, the "inhibitory" functions of the parasympathetic nervous system include increased activation of the digestive tract. And certainly "inhibition" has more than one meaning in psychology. The authors recognize that their hypotheses "span two quite disparate areas of function, that is, central autonomic versus intellectual functioning [pp. 29–30]," but they nevertheless fail to deal explicitly, not to say rigorously, with the problems this involves.

Thus although Broverman et al.'s argument proceeds through all the appropriate headings, the evidence for their repeated assertions of their thesis (first as hypothesis, then as conclusion at the beginning and end of almost every section) is inadequate to support their conclusion that a functional relationship exists between activation/inhibition or adrenergic/cholinergic neural processes or sympathetic/parasympathetic activity and the kinds of cognitive tasks listed in Tables 1 and 2 or described in the text as perceptual-motor and perceptual-restructuring.

RELATION OF ANDROGENS AND ESTROGENS TO THE "ACTIVATION–INHIBITION" BALANCE

This final phase of Broverman et al.'s argument seeks to establish a connection between estrogens and androgens, sympathetic/parasympathetic balance, and cognitive performance. They cite evidence from "rats, the sow, and, possibly, the human [p. 37]" indicating that estrogens are related to "activity" in the sense of spontaneous motor activity, and that androgens are also related to spontaneous activity (in rats and goldfish), but to a lesser extent than estrogens. This sort of evidence of the effects of hormones on the "behavioral expression of the activation–inhibition balance [p. 39]" leads them to think that "the cognitive activation–inhibition differences between the sexes in both animals and humans, therefore, may be primarily due to sex differences in estrogens [p. 39]." (As passages such as these illustrate, it is not clear that the authors consistently use "activation–inhibition balance" as either a psychological construct or a neurological process. In fact, they stipulatively define activation and inhibition as having both behavioral and neurological referents, and their argument seems to rely heavily on a blurring of the distinction between the two.)

While the authors have not, at this point in their argument, actually equated "activation–inhibition balance" with sympathetic-parasympathetic balance, the paper requires a very careful reading in order to avoid the inference that they are the same. When the authors do turn directly to a consideration of possible mechanisms by which androgens and estrogens could affect the parasympathetic and sympathetic nervous systems, respectively, they suggest that the hormone effects are mediated through their influence on the enzymes which control the available stores of the two transmitter substances for the autonomic nervous system. Specifically, their proposal is that the enzyme monoamine oxidase "provides a rational mechanism for explaining the apparent activation effects of these ["sex"] hormones [p. 40]" since it both regulates the amount of adrenaline available as a transmitter substance in the sympathetic nervous system and is itself influenced by both androgens and estrogens. In the section on "Steroid 'Sex' Hormones and the Parasympathetic Nervous System," the authors cite as "perhaps the best evidence [p. 40]" of hormonal effects the fact that the concentration of choline acetylase, an enzyme involved in regulating the amount of acetylcholine available as a transmitter substance, rises in the hypothalamus of castrated rats. Apparently Broverman et al. are unaware of the fact that while adrenergic and sympathetic, cholinergic and parasympathetic are often used loosely as synonymous terms, the transmitter substance in the preganglionic (central) portion of the sympathetic nervous system is acetylcholine (Ruch & Fulton, 1960, p. 228). "Sex" hormone effects on choline acetylase concentrations in the hypothalamus, therefore, do not constitute evidence—indirect or otherwise—of hormonal effects on the *balance* between sympathetic/parasympathetic activity, which is the general point they are trying to establish. They offer no other evidence that "sex" hormones have any effect on the activity of the parasympathetic nervous system. (When Singer & Montgomery, 1969, commented on the nonidentity of cholinergic-parasympathetic and adrenergic-sympathetic, Boverman et al., 1969, replied that these were "arbitrary problems of definition [p. 328]." This, of course, is true. In the context of the Broverman et al. paper, however, it is clear that that authors must be ignorant of the factual point that the sympathetic nervous system does involve cholinergic transmission, since they have cited increased hypothalamic choline acetylase as evidence that sex hormones affect the parasympathetic nervous system. It is simply false for them to say that "No part of our argument needs to be altered, regardless of which terminology is employed [p. 329]." Their argument needs to be altered in light of the factual distinction embodied in a precise use of the terms sympathetic/parasympathetic and adrenegric/cholinergic.)

In the final section of their paper, Boverman et al. (1968) note that the "simple perceptual-motor versus inhibitory restructuring abilities need not be the only intrapsychic cognitive antagonism [p. 43]." With reference to the thesis (Piaget, 1950; Werner, 1957) that ontogenetic development, like phylogenetic evolution, involves integration and subordination of "primitive" mental processes to more "complex" ones, they suggest that the argument they have developed "could be applied to all stages of development [p. 43]." (One suspects that a more forthright exposition of the implications of this notion might be found in Funke (1968) where it was suggested that females constitute a "semi-human" missing link in the evolutionary passage from anthropoid ape to man (cf. human).)

In sum, Broverman et al. "feel that the general arguments of this paper may be brought to bear upon a much wider domain of behavioral phenomena . . . [p. 24]." Considering the nature and relevance of the evidence on which their arguments rest, such a generalization of their thesis seems at best premature. As it stands, their conclusion represents another in a series of purportedly objective, scientifically empty statements on the topic of sex differences. As Ellis (1934) has said,

> The history of opinion regarding the cerebral sexual difference forms a painful page in scientific annals. It is full of prejudices, assumptions, fallacies, overhasty generalizations. The unscientific have had a predilection for this subject; and men of science seem to have lost the scientific spirit when they approached the study of its seat [p. 119].

REFERENCES

ANASTASI, A. *Differential psychology.* (3rd ed.) New York: Macmillan, 1958.

BROVERMAN, D. M., KLAIBER, E. L., KOBAYSHI, Y., & VOGEL, W. Roles of activation and inhibition in sex differences in cognitive abilities. *Psychological Review*, 1968, **75**, 23–50.

BROVERMAN, D. M., KLAIBER, E. L., KOBAYSHI, Y., & VOGEL, W. Reply to the "comment" by Singer and Montgomery on "Roles of activation and inhibition in sex differences in cognitive abilities." *Psychological Review*, 1969, **76**, 328–331.

ELLIS, H. *Man and woman.* (8th ed.) London: A. & C. Black, 1934. Cited in A. Montagu, *The natural superiority of women.* (Rev. ed.) London: Collier-Macmillan Ltd., 1968. P. 62.

FUNKE, M. *Are women human?* Cited in A. Montagu, *The natural superiority of women.* (Rev. ed.) London: Collier-Macmillan Ltd., 1968. P. 62.

GARAI, J. E., & SCHEINFELD A. Sex differences in mental and behavioral traits. *Genetic Psychology Monographs*, 1968, **77** (2), 169–299.

MACCOBY, E. E. Sex differences in intellectual functioning. In E. E. Maccoby (Ed.), *The developement of sex differences.* Stanford: Stanford University Press, 1966.

OETZEL, R. M. Annotated bibliography. In E. E. Maccoby (Ed.), *The development of sex differences.* Stanford: Stanford University Press, 1966.

PIAGET, J. *The psychology of intelligence.* New York: Harcourt, Brace, 1950.

RUCH, T. C., & FULTON, J. F. (Eds.) *Medical physiology and biophysics.* Philadelphia: W. B. Saunders, 1960.

SINGER, G., & MONTGOMERY, R. B. Comment on roles of activation and inhibition in sex differences in cognitive abilities. *Psychological Review*, 1969, **76**, 325–327.

TYLER, L. E. *The psychology of human differences.* (3rd ed.) New York: Appleton-Century-Crofts, 1965.

WERNER, H. *Comparative psychology of mental development.* New York: International Universities Press, 1957.

WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception.* New York: Harper, 1954.

WITKIN, H. A., DYK, R. B., FATERSON, H., GOODENOUGH, D. R., & KARP, S. A. *Psychological differentiation.* New York: Wiley, 1962.

(Received July 29, 1971)

New
From the Viking Press/Esalen Institute publishing program

A.H.MASLOW
THE FARTHER REACHES OF HUMAN NATURE

This first posthumous work serves as an extension of Maslow's classic *Toward a Psychology of Being*. It is a wide-ranging synthesis of his inspiring and influential ideas on biology, synergy, creativity, cognition, the hierarchy of needs, and the role of science in the expanding study of human nature. His theories of transcendence—the highest and most inclusive levels of human consciousness—and the evidence of "peak experiences" are developed in absorbing detail. Essays, notes, talks, reports, and memoranda, as he put them together here, combine to give a personal and highly informative and exciting look at Maslow and his work. $12.50

SECRETS of the TRADE
Madness, Creativity, and Ideology
by Joe Kennedy Adams

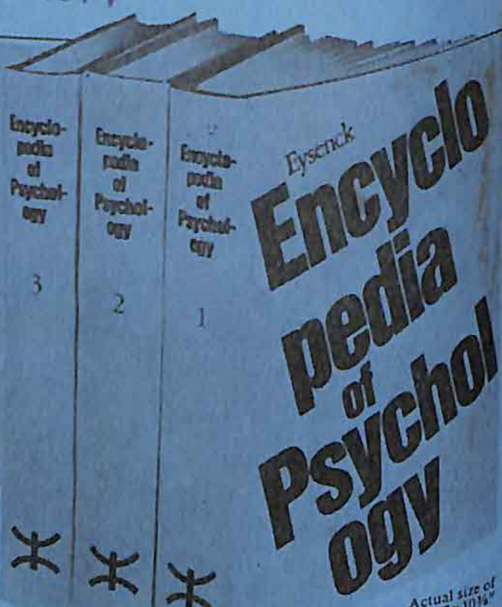
Dr. Adams talks about the creative mind, and its possible involvement with mental illness. He speaks personally as well as academically, since much of his material was gathered firsthand during his bouts in a sanitarium. Convinced that "many thousands of people are incarcerated unnecessarily, that millions are suffering unnecessarily, and that enormous amounts of creative energy are being wasted and destroyed instead of being utilized in a constructive way," the author's provocative book clearly questions the "experts' " knowledge of just where the fine line between madness and mad genius should be drawn. $12.50

For a complete descriptive listing of all Viking books in psychology, write:
Dept. ATE—PR, The Viking Press,
625 Madison Avenue, New York, New York 10022

2 FEB 1974

Now, in one three-volume encyclopedia...

The first comprehensive standard reference work dealing with all facets of psychology today



Actual size of each volume, 7 x 10⅛"

Here is the most accurate, readable compendium of current psychological thought ever assembled, a practical tool for professionals and laymen alike—the new, totally up-to-date ENCYCLOPEDIA OF PSYCHOLOGY.

For every man and woman concerned with this fascinating and vital discipline, many advances in the field of international psychology have been a source of frustration as well as excitement. New experiments and insights have undeniably changed the course of man's journey toward self-understanding. But for most readers, the results of current research and thought are inaccessible—available only in specialized publications, and often in a foreign language.

Now, in response to the need for a single comprehensive reference work in the field of psychology, Professor H. J. Eysenck has collaborated with hundreds of international specialists to present the three-volume ENCYCLOPEDIA OF PSYCHOLOGY — a major tool for the understanding of current psychological knowledge.

COMPLETE, UP-TO-DATE, READABLE

The new ENCYCLOPEDIA OF PSYCHOLOGY contains over 1200 pages and more than 120,000 words, with 5000 entries and 300 major articles by over 300 distinguished contributors. Each major entry offers a full definition of terms, an historical overview, a rundown of methodology, discussion of research and scientific controversies, and descriptions of the varying approaches of different psychological schools and related disciplines, as well as a complete bibliography. The ENCYCLOPEDIA OF PSYCHOLOGY provides both essential background information and a summary of leading international opinion on current issues. In addition, a thorough system of cross-references enables the reader to see the key concepts of psychology in a meaningful context.

INTERNATIONAL REFERENCE WORK

The ENCYCLOPEDIA OF PSYCHOLOGY is the first venture of its kind in this field. It is being published in five languages simultaneously, and internationally known authors from twenty nations have participated in its preparation. Contributors include Anne Anastasi from the United States on *Differential Psychology;* A. R. Luria from the Soviet Union on *Brain Pathology;* J. D. Teasdale from Great Britain on *Drug Dependence;* L. Shaw from South Africa on *Accident Proneness;* F. Klix from East Germany on *Cybernetics and Psychology;* and P. E. Vernon (Canada) and T. Takuma (Japan) on *Personality*. Many additional contributions have been made by scientists from countries such as Czechoslovakia, Italy, Israel, and Japan, whose studies until now have been all but completely inaccessible to American readers. In Professor Eysenck's words, "It is our hope that this Encyclopedia will play a part in making our science more truly international than it has been, by drawing attention to work done in other countries, and making people more aware of the international character of modern psychology."

By subscribing now to the ENCYCLOPEDIA OF PSYCHOLOGY, you may assure yourself of receiving special pre-publication prices with substantial savings. Send your order today on the convenient coupon.

300 AUTHORITIES FROM 20 COUNTRIES:

UNITED STATES: L. Berkowitz on *Aggression*. A. Bandura on *Socialization*. J. E. Birren on *Aging and Gerontology*. R. B. Cattell on *Factor Analysis*. E. R. Hilgard on *Behaviorism*. ☐ **GREAT BRITAIN:** R. Lynn on *Arousal*. Irene Martin on *Emotion and Stress*. A. E. Maxwell on *Correlational Techniques*. G. R. Wilson on *Objective Tests*. I. Oswald on *Sleep*. G. S. Spinks on *Psychology of Religion*. S. Rachman on *Aversion Therapy*. ☐ **SOVIET UNION:** A. R. Luria on *Aphasia*. A. N. Sokolov on *Thought and Inner Speech*. ☐ **FRANCE:** J. Leplat on *Occupational Psychology*. ☐ **HUNGARY:** F. Illyes and A. E. Lanyes on *Educational Guidance*. ☐ **GERMANY:** W. Schraml on *Clinical Psychology*. L. Pongratz on *Mind-Body Problem*. H. Kirchhoff on *Expression*. ☐ **CANADA:** D. E. Berlyne on *Experimental Aesthetics*. A. B. Kristofferson on *Attention. And many many more.*

H. J. Eysenck, Editor of the ENCYCLOPEDIA OF PSYCHOLOGY is Professor of Psychology at London University, and Director of the Institute of Psychiatry at Maudsley Hospital. He has been hailed by *The London Times* as "one of the greatest living psychological experimenters."

72-7

HERDER AND HERDER
A Division of McGraw-Hill Book Company
232 Madison Ave., New York 10016

Please send__________set(s) of the ENCYCLOPEDIA OF PSYCHOLOGY in three volumes at one of the special pre-publication subscription prices being offered for a limited time.

☐ M41073 M41074 M41075
Pre-publication subscription price: $67.50. We will bill you $22.50 (plus small mailing charge) for each volume you receive.

☐ M41070 M41071 M41072
Pre-publication subscription price: $64.50 ($21.50 per volume). We will bill you $64.50 (plus mailing charge) for the complete three-volume set when you receive the first volume. ☐ Additional savings: There will be no postage charge if your check accompanies your order.

Name__________

____

Address__________

__________ Zip__________

Payment $__________ Enclosed ____ Please Bill

Under each purchase plan I will receive Volume I now - Volume II in May, and Volume III in September.

[illegible]

1/24/74

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 79, No. 3 May 1972

Articles

Nonverbal Behavior and Nonverbal Communication
Morton Wiener, Shannon Devoe, Stuart Runbinow, and Jesse Geller 185

Ontogeny of Memory
Byron A. Campbell and Norman E. Spear 215

Mechanisms of Habituation in the Brain Stem
Philip M. Groves and Gary S. Lynch 237

Conjoint-Measurement Framework for the Study of Probabilistic Information Processing
Thomas S. Wallsten 245

Theoretical Notes

On Testing the Independence of Associations
Douglas L. Hintzman 261

Independence of Associations Tested: A Reply to D. L. Hintzman........**Edwin Martin and James G. Greeno** 265

A Stochastic Model for Free Association Response Hierarchies?..................**Vaira Vikis-Freibergs** 268

On the Distinction between Density and Crowding: Some Implications for Future Research
Daniel Stokols 275

A Comment on D. J. Moore and D. A. Shiek's "Toward a Theory of Early Infantile Autism"....**Roger A. Webb** 278

A Response to R. A. Webb's Comments
Dewey J. Moore 280

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

DARYL BEM
HARLEY A. BERNBACH
PAUL EKMAN
WILLIAM K. ESTES
J. L. FREEDMAN
R. GELMAN
ELEANOR J. GIBSON
B. F. GREEN
RALPH N. HABER
N. JOHNSON

LEON J. KAMIN
FREDERICK H. KANFER
HOWARD H. KENDLER
TRACY KENDLER
PAUL A. KOLERS
DAVID H. KRANTZ
MARVIN LEVINE
R. DUNCAN LUCE
GEORGE F. MAHL
E. MARTIN

JAMES MCGAUGH
RONALD MELZACK
DANIEL OSHERSON
SEYMOUR PAPERT
M. PETERSON
P. M. A. RABBITT
ROBERT A. RESCORLA
R. L. SOLOMON
R. SOMMER
JOHN E. R. STADDON

N. S. SUTHERLAND
D. L. THISTLETHW
THOMAS TRABASSO
MILTON A. TRAPO
HARRY C. TRIAND
BENTON J. UNDER
GEORGE WOLFORD
J. I. YELLOT, JR.
JOSEPH L. YOUNG

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts sub to the *Psychological Review*. In each issue the Editor will list in the *Review* those editorial consultan have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowle

The *Psychological Review* is devoted to articles of theoretical significance to any area of sci endeavor in psychology. Except for occasional articles, manuscripts exceeding 7,500 words w discouraged. Ordinarily, manuscripts which consist primarily of original reports of research or of re of the literature should be submitted to other journals. Methodological and statistical articles ar appropriate for the *Psychological Review*. See the editorials in the *Review* (1965, **72**, 1–2, 1971, **78**, for further details.

Beginning with the 1971 issue, Letters to the Editor will be published on an experimental basis. Le not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate she paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obta from the APA Central Office.

All manuscripts must be submitted in triplicate. In the reference list, give journal titles in full; do abbreviate. Figures should be submitted as glossy prints; duplicate figures may be photographic or pe drawn copies. Authors should retain a copy of the manuscript in case there is loss in the mail.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of sev months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own artic the appearance of articles by other contributors is not thereby delayed.

Manuscripts should be addressed to the Editor, George Mandler, Department of Experimental P chology, South Parks Road, Oxford OX1 3PS, England.

Copyrighted material. The following policies govern reprinting or other forms of reproduction materials copyrighted by the American Psychological Association: Written permission must be obtain for copying or reprinting any tables, figures, or text of more than 500 words in length. Persons or orga zations who obtain permission to reproduce any such copyrighted materials must include on the first page the reproduced materials the exact copyright notice which appeared on the Association's original publicatio There will be a charge of $10.00 per page for such permission except when the requester is the author of t article to be reprinted, or when the article is to be reproduced in a limited number of copies solely for i structional purposes.

HAROLD P. VAN COTT
Managing Editor

LORRAINE BIROS
Technical
Editor

HENRY S. SALMON
Advertising and Business
Manager

ANNE REDMAN
Subscription Manager

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1200 Seventeenth St., N. W., Washington, D. C. 20036

$10.00 volume

$2.00 issu

Second-class postage paid at Lancaster, Pa.

Send all communications, including address changes, to 1200 Seventeenth St., N. W., Washington, D. C. 20036. Addre changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from addre changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postag Other claims for undelivered copies must be made within four months of publication.

PSYCHOLOGICAL
REVIEW

Diary No. 727(8)
Date 42
No. Lib
Bureau Edu'l Pay Research.

VOL. 79, NO. 3

MAY 1972

NONVERBAL BEHAVIOR AND NONVERBAL COMMUNICATION[1]

MORTON WIENER,[2] SHANNON DEVOE, STUART RUBINOW, AND JESSE GELLER[3]

Clark University

In this paper an attempt is made (*a*) to articulate some of the suppositions implicit in different approaches used to study nonverbal behaviors as communication, (*b*) to reemphasize a conceptual distinction between nonverbal behaviors which can be considered as communications and other nonverbal behaviors, (*c*) to reformulate some issues in the study of such behaviors, (*d*) to establish definitions and criteria which can make possible some empirical investigations of nonverbal communication, and (*e*) to indicate how these definitions and criteria can be applied to the study of hand and arm movements as communicative gestures. It is hoped that the availability of the conceptualization, along with some means for operationalizing the concepts, will make it possible to initiate systematic empirical investigations of nonverbal communication in general.

In recent years there has been a marked resurgence of interest in behaviors variously called "body language," "expressive movement," and "nonverbal communication." The inclusion of the terms "language," "expression," and "communication" in these phrases suggests an act of transmission of information from one person to another; that is, the terms imply a subject who is actively making his experience public to some other person. The terms, as used, also suggest that there is a set of nonverbal behaviors which have shared meanings[4] in the same way that there is a set of verbal behaviors which have

---

[1] This study was supported in part by Public Health Service Research Grant 1 RO1 MH 17048-01 from the National Institute of Mental Health. Morton Wiener and Shannon Devoe have been continuously involved in the formulation of concepts and the writing of the many drafts of this paper from its inception to its conclusion, Stuart Rubinow has been instrumental in helping with the later reformulations and writing, Jesse Geller was involved in some of the early conceptualization and drafts and was primarily responsible for the detailed descriptions of the hand and arm movements discussed in the latter part of the paper. Merry Ann Jancovic and David Roth have been extremely helpful in editing and commenting on the later drafts. We also gratefully acknowledge Sue Novick Rhodes and Marc Kashinsky for their work in determining the reliability of the category system described here and Jill Rierdan and Leonard Cirillo for their helpful comments.

[2] Requests for reprints should be sent to Morton Wiener, Department of Psychology, Clark University, Worcester, Massachusetts 01610.

[3] Now at Yale University, Department of Psychiatry, New Haven, Connecticut.

[4] It is important to note that we will use the terms "code," "referent," and "meaning" without implying a particular bias for one or another theory of meaning or theory to account for the ways in which the code component and its referent are related. While the way we use the terms may seem to imply a preference for a referential theory of meaning, we are not committed to this view (see Alston, 1964, for his categorization of theories of meaning). As far as we can determine, the issues raised here about code usage seem to apply equally to all theories of language. Fundamentally, we are concerned about the problems involved in trying to "discover" a "language" rather than to explain how it is possible to use language.

© 1972 by the American Psychological Association, Inc.

shared meanings. It would be reasonable, then, to expect that investigations conducted *under the rubric of nonverbal communication* would be concerned with the transmitting subject (encoder) and with the behaviors (code) by which he makes his experience public to another (decoder). Instead, a review of the extensive literature in this area indicates that most investigations seem to be concerned with the significance which some observer can attribute to a particular behavior; that is, the emphasis seems to be primarily on decoding. Even given a decoding approach, there is little consensus about any set of behaviors which can be considered to serve communication functions, and the literature consists, for the most part, of a fragmented and unsystematic array of reports, with almost any conceivable behavior considered by one or another investigator to have some communicative significance. Thus, such diverse movements as body or head positions, lint picking, foot kicking, scratching, gross postural shifts, and hand and arm movements such as palm up during a verbalization have been considered equally cogent and relevant for an investigation of communication.

Whatever the conceptual approach (e.g., transactional, psychoanalytic) or the concerns of the investigators (e.g., behaviors of individuals or of groups), most if not all investigators seem to take a decoding perspective. Investigators who share this perspective also appear to share an assumption that if the observer can make an inference about an individual from his behavior, then the behavior can be considered to be a communication. Unfortunately, this kind of implicit assumption seems to fuse the notion of sign with the notion of communication.[5] For us, "sign" implies only an observer making an inference from, or assigning some significance to, an event or behavior, while "communication" implies (*a*) a socially shared signal system, that is, a code, (*b*) an encoder who makes something public via that code, and (*c*) a decoder who responds systematically to that code.[6]

When investigators have not clearly distinguished sign from communication, the statement, "I take *x* to be a sign of *y*" becomes transformed into, and considered equivalent to, the statement, "He communicates *y* via *x*." The logical difficulties in this transformation can easily be seen when it is applied to inanimate events. For example, the fact that one may take dark cumulus clouds to be a sign of impending rain does not mean that the clouds communicate, *via a code* which includes darkness and cumulusness, that it will soon rain. Nevertheless, investigators have frequently transformed statements such as "I take behavior *x* to be a sign that he is angry" into statements like "He communicates his anger via behavior *x*." It is clear that while any and all behaviors of an individual or group may be used by an observer as a basis for making inferences (i.e., any behavior may be taken as a sign of something else), communication implies something about the activity of the behaving one and the specific significance of the act.[7]

Movements are Expressive

This kind of confusion among signs, inferences, and communications can better be exemplified by showing two ways in which the phrase "movements are expressive" can be interpreted. One interpretation takes the phrase to mean: "In principle, under specified conditions, everything a person does, every action or movement he makes, can be used by a trained, knowledgeable, and sensitive observer as a basis (either from the behaviors themselves or from inferences based on these behaviors) for categorizing (or inferring about) the individual—his state, thoughts, feelings, personality, or cultural patterns." In this context, the phrase "everything a person does . . . can be used as a basis . . . for inferring about . . ." can be abstracted to the statement, "If behavior x occurs, an observer may infer a concomitant state y, or an antecedent z, or both, as *correlates* of the behavior x."

---

[5] The issues discussed in this paper are relevant primarily to those investigators, whatever their concern or their philosophical or theoretical orientation, whose definitions distinguish, either implicitly or explicitly, communication or language behaviors from other behaviors. For those investigators who make no such distinctions, for example, those who treat *all* behaviors as signs, some of the issues discussed in this paper may not be relevant (see Footnote 7 also).

[6] By "code" we will mean only a shared set of behaviors which have shared referents, as in Morse code. We will use the term "system" in this context, at least, to include the code and the principles for usage of code elements. For us, as for others, "language" additionally implies some rules governing organization and combination of code elements (i.e., grammar, syntax).

[7] While we will go to considerable length to distinguish sign and communication behavior, this distinction is important only if the concern of the investigator is with communication per se. If the concern is with understanding the behavior of a particular individual, signs and communications may be equally valuable data sources, and the distinction between these two sets is not very relevant. If, however, the concern is with communication—that is, a system independent of particular individuals or particular referents—then these distinctions are important.

A second interpretation of the phrase "movements are expressive" takes "expressive" to mean that a behavior is a manifestation of something which is itself not observable. The apparent reasoning in this interpretation is that since an observer can infer something (e.g., a genotypical characteristic, psychodynamic conflict, or group membership) about the speaker from a movement—much as he can from a spoken word—then what has become evident to the observer must have been *made* manifest by the actor. The movement is then considered to have the characteristics of a communication; that is, something is being made public by the doer. This interpretation of "movements are expressive" then involves a shift from an observer making an inference, given certain behavioral events, to a subject who is manifesting something, that is, communicating via these behaviors. In this interpretation, there is yet a further transformation. What is being made evident is known by the observer who makes the inference, while the doer apparently neither knows what is being made manifest (his assumed communication) nor does he make any inference about his movement. For this reason, the phrase "movements are expressive" is further transformed to read: "A person, in doing x, is revealing y covertly." In this form, the behavior is seen as a manifestation of something covert which is being involuntarily communicated or revealed.[8]

For investigators who use this second interpretation of the phrase, if something is covert and being revealed, it seems reasonable to infer that what is being revealed must be unacceptable or it would be expressed in a more overt form. Further, these investigators reason that if something is covert, then its behavioral manifestation must be in a form which should not be recognizable by the actor, that is, symbolic in the psychoanalytic sense. In final form, the phrase "movements are expressive," as interpreted from this perspective, seems to mean: "A person's behavior x communicates in some symbolic form his unacceptable thoughts, affect, impulses, etc."

The first interpretation of "movements are expressive" was concerned with an observer who makes inferences about traits of the subject, about concurrent cognitive or affective experiences of the subject, and/or about antecedents of the behavior. In contrast, the transformations of the second interpretation have subtly shifted the focus to a subject who intentionally or unintentionally symbolizes unacceptable content by means of these behaviors. In these transformations there is also a shift from an observer who is knowledgeable about a set of categories, uses them reliably, and has a set of correlated events with which he associates the categories in his observation, to an observer who has been specially trained to interpret symbolic forms of expression within the context of some particular theory (e.g., psychoanalytic).

While the transformation patterns described above may seem like word games, an examination of the literature on body language, expressive movements, and nonverbal communication would indicate that all too often an observer's inference seems to be interpreted as a subject's communication, either explicitly by the investigator himself or by some ambiguity in the conceptualization which makes it possible for a less sophisticated reader to make the transformation. The shift from an inference-making observer to a communicating subject takes different forms, depending on the behaviors being investigated (e.g., idiosyncratic vs. socially shared) or the theoretical assumptions of the investigator (e.g., psychoanalytic theory vs. social learning or role theory). We will attempt to exemplify the transformation from sign to communication as it occurs in the context of a number of representative approaches to the study of nonverbal communication by referring to investigators who have been trying to understand the individual and to investigators concerned with group behavior.

---

[8] It is all too tempting to infer unconscious communication from movements, since the subject often seems to be neither as aware of his movements nor as able to reconstruct them as he is aware of and able to reconstruct his spoken words. The issues of intentionality and awareness are discussed later.

In selecting our exemplars, we have tried whenever possible to select the most influential investigators, many of whom have also been concerned with the very transformation issue with which we are concerned. However, we will attempt to demonstrate that despite the concern for distinguishing communicative from noncommunicative nonverbal behavior, the decoding perspective which has been the primary focus of investigations in this area can be seen as a main source of confusion in the field in general. Simultaneously we hope to explore within each example additional issues which we feel are important if we are to arrive at a conceptual framework which will permit some systematic investigations of nonverbal communication from an encoding perspective. In any case, none of our criticisms, either implicit or explicit, should be taken as minimizing the importance of the contributions of the investigators whose works we are analyzing, for without their pioneering work there would be nothing to extend.

Decoding Perspectives with Focus on the Individual

The writings of many of the psychoanalytically oriented investigators clearly show the way in which the transformation from movement as sign to movement as communication could occur.[9] Given the assumptions and the context within which psychoanalytically oriented investigators have traditionally worked, it is seductively easy to make such a transformation. The patient's task in the psychoanalytically oriented situation is to make his past and present experience public via verbal language; the analyst's task is to understand or "make sense of" the patient's report. To understand, in this tradition, is to seek indexes of content and experiences which are not being verbalized. Thus, with verbal material it is the latent rather than the manifest content which is taken to be most critical. This concern with latent content goes with the assumptions that everything which is available to consciousness should be communicable through consensually shared verbal forms, and that anything which is not available to consciousness (i.e., anything being defended against) will be evident only in a form in which the patient himself could not recognize the content—that is, the material is in symbolic form.

From our viewpoint, the latent content may be seen as an inference drawn by the therapist based on the patient's verbalizations. Although we agree that the manifest content is a communication, when verbalizations are interpreted within a psychoanalytically oriented context, they seem to be used as behaviors, with the behaviors treated as signs of some other particular state, affect, or experience. What we would take to be an inference seems to be taken by these investigators to be some latent content. It would seem, however, that in this psychoanalytically oriented view, both manifest and latent content are often treated as equivalent in terms of communications. For both types of content, many psychoanalytically oriented observers seem to see themselves as decoding a communication encoded by and emanating from the patient. Indeed, to talk about latent content at all seems to presuppose that those meanings are "in" the utterance and, by extension, "in" the speaker, rather than that those meanings are the listener's inferences. The transformations thus can move subtly from "I can infer

---

[9] There are clearly a number of psychoanalytically oriented investigators who have been explicit in separating their inferences from the patient's communications (e.g., Mahl, 1968).

x from his utterance" to "his utterance communicates x" to "he communicates x symbolically and thus unconsciously." To the extent that such transformations occur for verbal material for any psychoanalytically oriented investigators, the same transformations could occur even more easily when dealing with nonverbal behaviors, especially if the nonverbal behaviors are taken as prima facie manifestations of some latent content that is by definition unable to be verbalized. The patient can then be seen as using nonverbal behaviors unconsciously to communicate that which he cannot or will not communicate verbally. Deutsch (1959) made explicit this kind of assumption: "Postural behavior is already a stage of transformed instinctual drives, a kind of acting out in a rudimentary way of expression that could not reach the verbal level [p. 129]." Feldman (1959), Reik (1954), and Robinson (1965) seem to have taken positions very similar to this one.[10]

There are a number of interrelated problems for understanding nonverbal communication as a system which arise if nonverbal behavior is considered to be unconscious encoding of experience. First, if the unconscious encoding is considered to be idiosyncratic (i.e., if each person encodes his experience in a different way), then there seems to be little if any possiblity of deriving a socially shared system for encoding experience in movements. If a consensually shared system of encoding experience is taken to be one of the criteria of a communication system (as in verbal language), idiosyncratic unconscious encoding would not meet this criterion. From a research concern, the view that each individual encodes his experience in some idiosyncratic behavior could only yield a series of case studies showing how a particularly insightful and experienced observer has cracked a particular idiosyncratic code. Such intensive case studies are common in many psychoanalytically oriented contributions to the literature on nonverbal communication. However, there appear to be no principles determining which behaviors are selected for study or interpretation by the investigator, and there seems to be little overlap of research efforts on the significance of particular movements. It is unclear how any number of case studies could increase our knowledge of nonverbal communication as a system.

Second, if unconscious encoding of experience is assumed to take consensual rather than idiosyncratic forms, then this assumption would at the least be more consistent with most views about communication. However, this view requires some consideration of the status of the decoders. It could be assumed that there is a select group of decoders, that is, decoders (e.g., psychoanalysts) who know a code that others do not know. This assumption would result in an unusual distribution of encoders and decoders. In most communication groups, there seem to be more decoders than encoders of the communication system. For example, children decode verbal language forms which they do not use in encoding, and beginning students of a foreign language can often decode more than they can encode. Thus, the notion of a communication system which contains all members of the group as encoders, but only a few as decoders, seems contrary to what we find about encoding and decoding within other communication systems (e.g., verbal language). Additionally, the assumption of consensual unconscious encoding requires some empirical documentation of the posited significance of a movement. Further, if every movement were taken to be an unconscious consensual encoding of experience, any research effort would leave most movements unstudied, and no finite number of experimental studies would justify logically the generalization that all movements are universally encoded communications. If only some movements are to be taken to be encodings, it is difficult to know the basis for determining which movements are relevant for study.

---

[10] Certain procedures occurring in psychoanalytically oriented psychotherapy seem almost guaranteed to elicit material which is consistent with psychoanalytic assumptions about manifest and latent content, and which thus precludes questioning of the assumption of unconscious communication. Therapists frequently note particular behaviors of their patients, either from direct observation or from the patient's report of past behavior. To understand the patient better, the therapist will then ask him to talk about his experience concurrent with that behavior. For example, a patient may be directed to "tell me how you felt (what you were thinking, etc.) when you did x (some particular behavior)." Most patients will respond to this directive with some verbal communication about some experience. It is tempting then to assume that the nonverbal behavior was a communication of what was later communicated in consensual verbal forms. It is even easier to maintain this assumption if the therapist has made an inference about the patient's experience from the behavior, and this inference is supported by the patient's subsequent verbalizations. If verbalization does not occur, the therapist may still infer resistance or repression, thus supporting his assumption that the behavior was a nonverbal communication.

It is possible, of course, to assume that all behaviors are consensually unconsciously decoded as well as consensually and unconsciously encoded, with only a few members of the communication group able to decode consciously what others encode and decode unconsciously. If this view of nonverbal behavior as communication is taken, it is incumbent upon the conscious decoders to specify at the very least the explicit criteria they use for determining the consensual referent of particular behaviors. To our knowledge, no attempt at such specification of criteria has been made.

In sum, the kind of approach to communication exemplified by many psychoanalytically oriented investigators is so broad that any and all behaviors are considered to be communications. It is difficult to see in this approach any basis for developing a systematic study of nonverbal behaviors as communication given (*a*) some of the assertions in the theory and the procedures of those who use the theory in this way, and (*b*) the fact that assigning of a behavior to the category of communication seems to be based solely on the observer's ability to make an inference from the behavior. If no distinction is made between signs and communications, it is unclear how such an approach could contribute anything to the study of communication as a special system or as a special set of behaviors. In fact, there seems to be no logical basis for excluding from communication instances any occurrence about which the observer might make an inference. This failure to distinguish sign from communication, and inference making from decoding, would yield a fragmented and unsystematic research literature, whose contribution to the study of communication as a special system is uncertain. In our view, it would seem that an attempt must be made to distinguish sign from communication conceptually before communication can be investigated empirically in any systematic way.

There is a group of investigators (e.g., Dittmann, 1962; Duncan, 1969; Sainsbury, 1955) who do not seem to be concerned about making the distinction between movements as signs and movements as communications. In contrast to the psychoanalytically oriented investigators discussed above who seem to treat all behaviors as communications, the latter investigators appear to treat all behaviors as signs. For example, Sainsbury correlated the occurrence of movement in any part of the body with indexes of affect (e.g., verbalization, galvanic skin response); Dittmann correlated the occurrence of head, arm, and foot movements with verbal indexes of affect; and Duncan (personal communication, 1969) correlated movements from various parts of the body with each other and with verbalizations. However, since these investigators also use participants in a therapeutic interview as subjects, the temptation is ever present for a reader to view the observed behaviors as communications of affect or other momentary states, much as these transformations are made by some psychoanalytically oriented investigators. Whatever the view of these investigators about the relevance or importance of making a sign–communication distinction, subsequent workers in the field appear to have taken these investigations as studies of communication, as evidenced by the inclusion of such studies in the literature on nonverbal communication.[11]

---

[11] The studies cited above, which use correlation techniques and which imply that movements are expressive of other states, are similar to the earlier work of Allport and Vernon (1933) and Wolff (1935). These earlier investigators correlated movements and movement styles with other

Incidentally, while a correlational research strategy seems appropriate for the investigation of movements as signs or indexes of enduring or momentary states, this strategy seems less appropriate for a study of movements as communications. A communication system, by our definition at least, requires that (*a*) individual units of behavior be specified, (*b*) each of these behavior units have specific significance, and (*c*) for the most part, the significance of each unit be different from the significance of other units. In the correlative studies cited above, all movements within any one part of the body have been treated as interchangeable, with each movement thus having no special significance (e.g., Dittmann, 1962; S. Duncan, personal communication, 1969), or any movement of one part of the body has been treated as equivalent to any movement of any other part of the body (e.g., Sainsbury, 1955). It is difficult, then, to see how this kind of approach can be heuristic for the investigation of communication as a system, were anyone concerned about these data for the study of communication. The results of the all-behaviors-as-signs approach, like those of the all-behaviors-as-communications approach, have been either (*a*) a series of case studies in which particular movements are found to be correlated with particular indexes for a particular person in a particular context (e.g., Dittmann, 1962; S. Duncan, personal communication, 1969), or (*b*) a general conclusion such as movements in general express affect in general (e.g., Sainsbury, 1955). The case studies approach appears to yield findings which make it difficult to arrive at any statement of shared or concensual use of movements; the general statement, by its very all inclusiveness and lack of specificity, makes further research almost unnecessary. In sum, a movement-as-signs approach, however painstaking and time consuming, would not seem to lead to understanding of nonverbal communication, if communication is taken to be a socially shared system for the encoding of experience and its decoding by another member of the group.

Some investigators (e.g., Krout, 1935a, 1935b; Mahl, 1968) have distinguished conceptually between movements as signs and movements as communications. Krout distinguishes conventional gestures, which he defines as socially patterned and shared forms of behavior, from autistic gestures, which he defines as movements which are unique for a particular individual in a particular context. Mahl similarly distinguishes communicative gestures, which he defines as movements which everyone would agree are substitutable for a verbal utterance,[12] from autistic gestures, defined as movements which are not substitutable for a verbal utterance. In his conceptualizations, Mahl has included these autistic gestures in the class of informative movements. Despite these explicit attempts to distinguish between movements as signs and movements as communications, the use of a decoding approach in studies of gestures makes it difficult to maintain this distinction. As we have suggested, there does not seem to be any way of knowing from an observer's perspective whether one is decoding a communication encoded by the individual or whether one is drawing an inference from his behavior.

We hold that these difficulties are inherent in a decoding perspective, even for investigators who attempt to distinguish signs and communication, and can best be exemplified in some work of Krout (1935a, 1935b) and of Mahl (1968) with autistic gestures.

---

behaviors (e.g., handwriting) which in turn were viewed as indexes of a relatively enduring state (personality, character). The writings of these earlier investigators seem to have focused on individual differences and on the consistency of the behavior of an individual over time, and seem to have been less concerned with nonverbal communication. Nevertheless, these investigators' description of their subject matter as expressive movements seems to have prompted the inclusion of their work in the literature on nonverbal communication by some authors, and has influenced the research strategy of later investigators.

[12] At first glance, the terms "conventional gestures" and "communicative gestures" may appear to refer to the same class of events. However, Krout's category also includes a great many socially patterned movements (e.g., bread buttering, car driving) which do not appear to us to meet Mahl's criterion of substitutability for a verbal utterance.

Krout (1935a, 1935b), for example, states explicitly that "autistic gestures have no definite meaning, either for the subject or for the individual to whom the subject is responding [1935a, p. 401]." This statement seems to imply that autistic gestures are behaviors, neither encoded nor decoded. Krout further hypothesizes that these behaviors are explicit responses to internal stimulation, and he demonstrates through a variety of experiments that (*a*) a particular idiosyncratic behavior is related to a specifiable external stimulus (word), (*b*) the external stimulus is related to other external stimuli (words), and (*c*) from the exploration of the word association configurations an observer may infer the existence of an emotional conflict or complex within the subject. However, when Krout finds evidence of an emotional conflict/complex and comes to the conclusion that an autistic gesture is a disguised expression of other nonexpressible behaviors or of nonexpressible words (i.e., those behaviors or words whose overt manifestation has been inhibited), the reader is likely to associate these kinds of statements with unconscious expression of repressed material. The collapse of the sign–communication distinction for the reader is made even more probable when Krout (1939) states,

> Theoretically, autistic gestures occur because they give us temporary relief from tension and conflict. To gesture means to think frankly along certain lines, and to *express* our thoughts with impunity. By means of self-directed gestures, we can say things about ourselves which we are loath to regard as attempts at personal advertising [p. 172].

Although Krout at times states that his goal is interpretation of autistic gestures, when he later says that to gesture "means" or "expresses," he seems to imply that he has decoded the subject's self-communicative acts. In this way, the transformation from sign to communication seems to have occurred; if Krout himself has not made it, the unwary reader is likely to. Thus, Krout starts with a conceptual distinction between conventional gestures and autistic gestures which appears to be based on a distinction between behaviors as communications (socially shared patterns of encoding and decoding which serve to make experience public) and behaviors as signs (behaviors serving as a basis for inference). However, in coming to the conclusion that autistic gestures can be seen as possible substitutes for verbal utterances to oneself or others, he could be seen as attributing the inferences he makes about the behaviors to an encoding process on the part of the subject, blurring, at least for the reader, the very distinction between signs and communication which he had previously attempted to introduce.

Mahl (1968) clearly warns that an

> Interpersonal orientation to psychiatry . . . places one in the danger of viewing everything the patient does that is *informative* to the therapist or observer as though it were also *communicative* in motivation [p. 334].

However, his own use of terms and his analyses of behaviors make it difficult for the reader to maintain the distinction he so carefully posits between informative and communicative movements. In his own work with autistic movements, Mahl finds that inferences he has made on the basis of the behaviors are later validated by some verbal communication from the subject. He then seems to conclude that some autistic gestures can be seen as precursors of later verbalizations; that is, given the appropriate evidence, an autistic gesture can be interpreted as involving "spontaneous nonverbal anticipation of verbalization [p. 323]." Since Mahl's definition of communicative gestures specifies that the behavior can be substituted for a verbal utterance, and since his analysis of autistic behavior may well lead the reader to the conclusion that these behaviors are alternative forms of expression for that which is later expressed verbally, it would seem that the distinction between communicative and autistic gestures can be maintained by the reader only with great difficulty.

Whether or not there is a collapsing of the sign–communication distinction for Krout or for Mahl, the readers of these investigations have collapsed the distinction, as evidenced by the inclusion of these studies in the literature of nonverbal communication.

We hold that insofar as communication at the very least involves socially shared behavior patterns, the study of autistic behaviors seems nonrelevant to the study of nonverbal communication. To us, the very inclusion of studies of autistic movements in the literature on nonverbal communication is evidence that our concerns are not unwarranted. We note, however, that Mahl and Krout in their work on autistic gestures, and indeed the other approaches which have been discussed thus far, have been concerned with understanding a particular individual. With a focus on understanding an individual, anything the individual does or says can be used as a basis for inference about the person, and any distinction between sign and communication is an unnecessary nicety. When communication is the focus, the distinction is mandatory in our view (see Footnote 7).

While some investigators have been concerned with nonverbal communication primarily to make more sense of particular individuals, Ekman (1964, 1965a, 1965b) and his associates (e.g., Ekman & Friesen, 1968)[13] seem to have been more explicitly concerned with communication per se. Ekman has made a noteworthy attempt to focus explicitly and systematically on communication. Nevertheless, the early formulations offered by him (at least up to and including 1968) appear to include many of the conceptual and methodological problems discussed thus far. Like others who employ a sign approach to movements, Ekman and his associates employed a definition of communication which seems synonymous with the way in which we have used the concept of sign. Specifically, Ekman and Friesen (1968) state that

> the term *communicate* refers to the fact that observers are able reliably to decode information from viewing a sample of nonverbal behavior. There is no implication that the person enacting the nonverbal behavior intended to communicate nor any assumption that the communication is necessarily accurate [p. 186].

In this statement, a behavior is said to be communicative if a number of observers make the same inference on the basis of the behavior. Although behaviors such as gait as a sign of age, or hip sway as the basis for inference about sex, could logically be studied as communication within this definition, these investigators focus on making inferences about affect from films or photographs of subjects in interview situations.

The fact that such behaviors as gait and hip sway are not included in their behavioral samples suggests that there are some implicit assumptions about communication which may not be included in the more explicit definition. Ekman and his associates have sampled behaviors only in the special context of a psychiatric interview. In this particular interpersonal context (which many investigators of nonverbal behavior have used as a data source), verbal communication is taking place. It seems that movements which have shared meanings are more likely to occur in a verbal communication context than in the many other contexts (e.g., a ball game) which might have been sampled.[14] For Ekman and his associates, as well as for other investigators, the very choice of a verbal exchange as the context for sampling behavior has obscured some of the difficulties associated with defining communication on the basis of observer consensus. At the same time, it makes their particular definition and approach appear to be reasonable for the study of nonverbal communication.

Another problem highlighted in some of this earlier work of Ekman and his associates, but evident also in the work of many other investigators of nonverbal communication, is the assumption, implicit if not explicit, that nonverbal behavior in general is communication about affect, or somehow related to affect (e.g., defense against it). Assuming that nonverbal behaviors have significance only for communicating about affect appears to have limited to one domain the possible range of decoding or inference making by the observer. For example, in this earlier work Ekman and his associates used either a two-category system for inferring pleasant and unpleasant affective states (Ekman, 1965a, 1965b) or an adjective checklist of words such as "nervous," "fearful," "angry" (Ekman & Friesen, 1968). It is not clear on what grounds—logical, empirical, or otherwise—the significance possibilities of nonverbal communication have been limited to affect significance, while verbal communications are recognized as having the possibility of a variety of domains of significance (e.g., informational, indicative). This implicit constraint seems to have unnecessarily hindered the exploration of the possible communicative significance of nonverbal forms of behavior, even within a decoding definition of communication.

---

[13] The most recent work of Ekman and Friesen seems to involve a radical shift in perspective. We discuss this work in a later section of this paper.

[14] It is logically possible to hold that there is an inverse relationship of verbal to nonverbal communication instances (as in the hydraulic model), or that these two sets of behaviors are uncorrelated. However, the definition we will propose would lead us to expect most instances of nonverbal communication behavior to occur in social verbal communication contexts, rather than in social situations per se.

Another methodological point in these earlier works of Ekman, as well as in the work of others, deserves mention. In some earlier studies (e.g., Ekman, 1965a, 1965b) the visual stimuli which judges rated were selected from interviews by a time-sampling procedure. Such a procedure is consistent only with an approach which defines communication solely in terms of observer consensus (i.e., if any behavior on which a consensus can be reached is taken to be a communication, then any and all behaviors are equally relevant for study). From our viewpoint, however, this procedure is likely to sample both communicative and noncommunicative movements. It seems hardly possible, using this sample procedure and rating system, to differentiate (*a*) observers' *decodings* of communications encoded by the interviewee, and (*b*) observers' *inferences* from the interviewee's behavior.

One further difficulty in an approach such as that of Ekman and his associates involves the same kind of transformation from observers' inferences to subject's communication already noted for other investigators. In this earlier work, Ekman (1965b) explicitly disavows the possibility of studying encoding in communication, as in the following statement:

> "What the sender actually experiences, intentionally or nonintentionally, is not only almost impossible to determine, but is in no way necessarily equivalent to the indicative or communicative value of his nonverbal act [p. 393]."

However, there are several indications that Ekman and his associates, like most others, seem to have assumed implicitly that the subject has encoded what the observers inferred. For example, the very use of the word "decoding" implies that there is or has been encoding. Similarly, the use of such phrases as "the subject conveys" or "the subject transmits" also implies active encoding on the part of the subject. Further, the failure to obtain observer consensus is explained by invoking encoder variables, as in the following: "there might be limitations to nonverbal communication, either for this patient or in general [Ekman & Friesen, 1968, p. 169]." None of these terms (i.e., decoding) or statements (i.e., limitation of patient competence in nonverbal communication) appears consistent with the disavowal of concern with encoding on the part of the behaving subject. However, there is some indication that Ekman and Friesen are now more explicitly concerned with encoding in their study of nonverbal communication. In a recent paper (Ekman & Friesen, 1969), a category system for movements was offered which depended almost entirely on encoder variables as criteria for categorization. However, it is not clear how this latter approach will be made coordinate with their former work, nor how it will influence their ongoing research strategy. We recognize, with Ekman and Friesen, the difficulties involved in investigating encoding; we show how we propose to deal with these difficulties in a later section of this paper.

Decoding Perspectives with Focus on Shared Behaviors

In contrast to the investigators who focus on the behavior of individuals, another group of investigators included in the literature on nonverbal communication focuses on behaviors common to a group.[15] The behaviors of interest to the latter investigators are at least known to be shared; they are thus, from our standpoint, more likely to be acceptable as communicative instances than are behaviors specific to particular individuals. Nevertheless, we hold that it is difficult to determine which, if any, of these socially shared behavior patterns qualify as communications by our definition. Again, from our perspective, investigations of socially shared behavior, like the investigations of an individual's behaviors, have proceeded from the standpoint of the observer or decoder. Accordingly, the subtle transformation from observer's inference to subject's communication is as easy to make in investigating shared behaviors as it is in investigating individuals' behaviors. There would at first seem to be some advantage, for communications research, to focusing on socially shared behaviors. This advantage disappears if the research approach makes it difficult or impossible to explicate a basis for identifying a socially shared communication code, in contrast to learned social behaviors or rule-following behaviors.

While many investigators of shared behaviors seem to make the transformation from observer's inference to subject's communication, they differ somewhat from each other in terms of the number and kinds of behaviors they include in their study of communication. Some investigators seem to include in their investigations only those behaviors which occur in interpersonal contexts, while still others focus on behaviors co-occurring with verbal communication in an interpersonal context, with each successive limitation on the range of behaviors making the transformation seem less obvious to a reader. We attempt (*a*) to explicate how the transformation operates for some different investigators of these social behaviors, and (*b*) to consider some of the problems attendant upon each approach. In addition, we hope to make some further distinctions which we hold are necessary if we are to arrive at a definition of communication which will be consistent for both verbal and nonverbal behaviors, and if we are to be able to initiate systematic investigations of communication.

The work of Goffman (1959) is a good example of the way the transformation from sign to communication can occur (either for the investigator or for the reader) when the investigator seems to treat all socially patterned behavior as communication instances. Goffman points out that a number of behaviors can be reliably correlated with a person's position or role in the social system. He gives "insignia of office or rank; clothing, sex, age, and racial characteristics; size and looks; posture; speech patterns; facial expression; bodily gestures; and the like [Goffman, 1959, p. 24]," as examples of signs which can be used by an observer as a basis for inference about the person's membership in a particular group. Even type and placement of living room furniture can be included among the events which Goffman considers relevant for his concerns. When Goffman treats these behaviors and events as signs of group membership, there are no difficulties in accepting his insightful comments on the social scene. However, he often calls these same behaviors or events "sign vehicles for expressing" and "communications." The interchangeable use of the terms "sign," "sign vehicle," and "communication" may well lead a reader to believe that all sign behaviors are transmissions of information by the subject, rather than bases for inference by an observer. The transformation from sign to communication appears plausible in that (*a*) the correlates of the behaviors cited are often known by most members of the group, (*b*) members of the group sometimes report having taken these signs into account when interacting with another, (*c*) the behaviors whose correlates are known can be and sometimes are used by one person to manipulate another into attributing to the user these correlates of the behavior, and (*d*) Goffman uses the metaphor of the theater to demonstrate his use of signs.

---

[15] For this group of investigators, as well as for the others we discuss, we are concerned only with the possible confusion of the sign–communication issue in their writing. We recognize that the problem may sometimes be one for the reader rather than for the investigators themselves. Whether our criticisms are valid or not, at no point do we question the value of the insightful observations made by those investigators of the behavior of individuals or of particular groups in particular settings.

There are several difficulties which can be exemplified in this approach. First, while the correlates of some signs are generally known by members of the group[16] (e.g., the wearing of a mink coat is correlated with wealth), the correlates of other, equally frequently exhibited signs are not known except by a few expert observers. For example, some hair-grooming behaviors appear to be correlated with gender; males typically groom hair by running their fingers through it, while females are more likely to pat their hair into place. If the correlates of a behavior are unknown by the emitters of the behavior, it is difficult to see how the behavior (e.g., hair grooming) could be used to communicate the correlate (e.g., masculinity or femininity in this case) even though it is a sign for the expert observer. Second, even if the correlate of the behavior is known and the behavior is sometimes used, it would seem less than logical to conclude that the behavior is always a self-conscious enactment.

There is a more basic concern about an approach which involves sign behavior, but which implies enactment. For a sign to be predictive, a correlation between the sign event and the other attribute which is to be designated must already be evident; that is, the co-occurrence must be known prior to any individual's using the sign event. For example, it must first be noted that the "Jet Set" wear certain clothes before anyone new can use the wearing of such clothes to be identified as a "Jet Setter." We cannot distinguish, from the clothes wearing itself, whether the wearer is part of the original group or is using his knowledge of the correlation. Further, to hold that a *behavior learned by a group is learned so that the group membership of the individual will be identifiable* seems to us to distort what is meant by learned behavior sequences. We call someone a "Southerner" when he speaks with a particular pattern of intonation; the Southerner does not learn to speak with that particular pattern of intonation so that he will be identified as a "Southerner." It would seem strange to say that he enacts or uses a Southern accent rather than simply that he has learned to speak that way. Similarly, one group of girls may learn to wear particular kinds of clothes. When an observer notes some characteristic of this clothing which distinguishes it from the clothing of other girls, we can construct a classification in terms of his observation. This classification is introduced by the observer, not by the clothes wearers, even though the observer could now take the clothing as a predictor of subgroup membership. We would hold, then, that instances in which a sign is used for manipulation purposes are infrequent. Although some signs are sometimes used for manipulation or deception, it does not seem reasonable to treat all signs as if they were always used or enacted.

There is a second group of investigators (e.g., Birdwhistell, 1970) who also study socially patterned behaviors, particularly movements. Unlike some investigators who seem to fuse sign and communication, Birdwhistell has in some of his early writing explicitly distinguished expression from communication in terms of a sign–communication distinction.[17] From our perspective, one of the most important issues raised by this group of investigators is that they explictly hold that nonverbal communication could most fruitfully be considered and analyzed in the same way as are verbal communications. Given this commitment and focus, these investigators seem to select movements of all kinds and attempt to determine the basic units which can be considered the building blocks of all movements, much as the phoneme is considered the equivalent building block of words. In a sense, these investigators view their work as being analogous to the work of structural linguists who deal with verbal-vocal material. While investigators of verbal language and these investigators of movement share a common methodology, we can note an important difference between them. The structural linguists apply their analysis to verbal events which have already been accepted and designated as meaningful events. For the structural kinesicist, on the other hand, there is not, at least to our knowledge, a set of movements which have been designated as meaningful to which a structural analysis can be applied. If a structural analysis is to be applied to movements, then, it must either be applied to all movements, irrespective of their communicative value, or to those movements which the investigators have taken to be meaningful, in which case it seems likely that the selection of movements would be based on what the investigator can make sense of and therefore assume to be decoded. In any case, the criteria for selecting particular movements to analyze have not been spelled out by these investigators. Instead, their writings imply that the use of a structural analytic methodology will, in and of itself, lead to the discovery of which movements are meaningful and what the meaning of those movements might be.

---

[16] See later discussion under section "An Alternative Approach to Nonverbal Communication." The same logic which is offered to determine code use can be applied to the question of "Knowing" or "Known" by a respondent.

[17] This distinction is sometimes blurred in Birdwhistell's anecdotal reports of social behaviors.

We have no doubt that a structural analysis of events already designated as part of a communication system may be extremely helpful in explicating the principles for representing, encoding, and decoding within the system. However, it is unclear to us how a communication system can be specified via the use of a structural analysis. For example, while we agree that it is possible to analyze every vocal utterance into some combination of phonemes, we may still ask how it is possible to specify which combinations of phonemes are to be classified as morphemes in any particular communication system. Similarly, while we agree that it is possible to analyze any and all movements into some combination of kines, we are still in doubt as to the method or criteria by which it would be possible to conclude that the combination ABC is a kinemorph, and thus part of a particular movement communication system, while the kine combination ABD is not a kinemorph within that communication system. In sum, although we can understand how to proceed with a structural analysis after a communication system has been designated, we do not know how to discover a communication system by an analysis of the smallest common units of a modality.

Thus, if there is no a priori specification of the set of movements to be analyzed structurally, then Birdwhistell (1970) and others are left with the impossible task of analyzing every movement. Even were this task possible, it is still not clear how an analysis of structure could lead to specification of meaning of the occurrences unless it is assumed that structure and meaning are correlated in some as yet unspecified way. Taking verbal behaviors as the paradigm of communication, this latter assumption appears to be less than tenable.[18]

Birdwhistell (1970) seems to recognize the difficulty inherent in trying to determine the meaning of a movement solely on the basis of a structural analysis of the movements, since he notes that similar movements may serve different functions and different movements may serve similar functions. However, his solution to this dilemma seems to be a belief that the meaning of a movement will be discovered by a careful analysis of the total context(s) in which the movement occurs. Again, it is no clearer how either a structural analysis or a detailed specification of the context will be a basis for deriving meaning—unless the meaning of the context is already known or posited. Incidentally, Birdwhistell appears to compound somewhat the difficulties associated with this approach when he holds that the meaning of the movement will be known only when the *total* context is known. This latter task seems logically impossible. It is not clear how an investigator can know that he knows the total context. We hold with Wiener and Mehrabian (1968) that since we do not know whether there is anything we do not know, we cannot be certain we know the total. Empirically it would seem that this search for the total context would be endless, for it will always be possible to include something else in the total, as Birdwhistell seems to do when he now includes olfactory components and micromomentary muscle changes as part of his concern with the total context.

---

[18] In verbal communication systems, the relationship between symbols and referents appears to be quite arbitrary; that is, no consistent principles for relating symbol and referent are apparent. It seems unlikely that the relationship between symbol and referent in nonverbal communication systems would be any less arbitrary.

Somewhat tangentially, it seems to have become fashionable for some investigators of nonverbal communication to use instruments which allow much more detailed inspection of behavior than would be possible for the average observer of the behavior (e.g., Haggard & Isaacs, 1966). The use of such devices might be helpful for detailed description of the observed events, or even for drawing hypotheses of correlations between particular aspects of the behaviors and other ongoing events. It is difficult, however, to see how minute movements discovered with special equipment could be included as part of a communication system, if these movements could not even be observable by participants in ordinary interactions.

In sum, Birdwhistell's (1970) observations about socially shared patterns of movement have been most insightful, and his delineation of the particular muscles involved in specific constellations (e.g., smile) for particular subgroups has been informative. However, it is not evident how this type of structural approach will lead to the specification of the movements which are to be considered as communications rather than signs, nor how the meaning of any movement will be determined.

There are a number of investigators (e.g., Scheflen, 1964a, 1964b, 1965, 1967; Watzlawick, Beavin, & Jackson, 1967) whom we categorize as taking as nonverbal communication any behavior occurring in an interpersonal setting. While Scheflen sometimes asserts a distinction between communicative and informative movements, at other times he seems to be espousing a definition similar to Watzlawick et al., who explicitly define communication as "any behavior occurring in the presence of another." Investigators who hold this view seem to justify the use of the context (interpersonal setting) as defining communication behaviors, by trying to show how the behaviors which do occur in this setting can be understood by an observer. In showing how they make sense of the behaviors, they, like many others, seem to be making the transformation and fusing the notions of sign and communication which from our viewpoint require separation. For example, Scheflen (1965) shows how a particular behavior sequence can be understood as a within-family communication event: A mother leans forward to speak to a therapist; the father begins to tap his foot; the grandmother and the daughter (flanking mother) cross legs toward each other; the mother stops talking to the therapist and leans back. In an attempt to make sense of this recurring sequence, Scheflen supplies the following meanings of the behaviors: When the mother leans forward she is flirting with the therapist; when the father taps his foot he is signaling to grandmother and daughter to limit the mother's flirtatiousness; when the grandmother and the daughter cross their legs toward each other they are telling the mother to stop flirting; when the mother leans back she is saying that she understands their message.

From our standpoint, if these behaviors are communications they should be part of a code shared either by a cultural group or at the very least by members of this family. Since Scheflen (1965) does not report the occurrence of this behavior sequence in other families in therapy, it would seem to be, at best, a code which is used only by this family. However, other than the interpretation offered, no evidence is presented that code usage is involved. We hold that for the behaviors to be considered a code in our sense requires the possibility of (*a*) interchangeable use of code elements by different members of the family, and (*b*) occurrence of the same code elements in contexts other than the one reported. No evidence is presented that, for example, the grandmother taps her foot to tell the daughter to limit the mother's flirtatiousness, or that a leg cross in another context has the meaning of stop it. It is, therefore, difficult to accept the view that this particular behavior sequence does involve a code, even one shared only by this particular family. Further, without evidence that the emitters of the behaviors are encoding meanings via a shared code, it would seem that the meanings in the behaviors are inferences of the observer rather than that the behaviors are the subjects' communications.

In contrast to Scheflen's (1965) account, let us see how the same behavior sequence can be accounted for in terms of learned responses—an analysis in terms of cue–response learned sequences—rather than in terms of a special communication code. Mother leans forward when talking to a male; whenever mother talks to another male, father becomes anxious or restless; when he is anxious or restless, he taps his foot; when father taps his foot, grandmother and daughter recognize that he is upset (i.e., they recognize that in the past this sign behavior has been correlated with instances of upset); they are uncomfortable when father is upset; when they are uncomfortable they fidget, crossing legs, etc.; mother notes the moving around and takes the fidgeting and other signs of restlessness as discontent (or upset or whatever else it has correlated with in the past); she stops what she is doing, which makes it possible for the others to enter the interaction.

Since this and similar behavior sequences can be accounted for without reference to communication, it would seem unnecessary to account for these kinds of behaviors as communication. In all fairness to investigators such as Scheflen, the behaviors selected for analysis are those which ostensibly occur with some frequency and have some predictability of sequence in particular interpersonal contexts.

If we examine some of the apparent assumptions underlying the definition and selection of instances for study, we may begin to understand the general acceptance of this type of approach. Verbal behaviors can elicit feedback from others, the patterns of verbal exchanges are somewhat predictable, and verbal behaviors involve the communication of meanings. From the perspective of this group of investigators, the nonverbal behaviors in interpersonal contexts which they select also seem to elicit responses from others, and these nonverbal behavior sequences also seem somewhat predictable; the nonverbal behaviors are then apparently assumed to have the same status as verbal communications. If this logic is reduced to syllogistic form, it would read approximately: All A (the nonverbal behaviors in interpersonal contexts) are B (predictable, feedback eliciting, etc.); all C (communications) are B (predictable, feedback eliciting, etc.). The conclusion that all A are C is clearly untenable. The apparent logic in an approach that includes all behaviors in interpersonal contexts as communications suggests that the justification of this kind of definition is on tenuous grounds. We recognize that definitions are arbitrary—subject only to agreement and not to proof. However, if all behaviors which could occur in the presence of another person are to be considered to be communications, what then is *not* communication? If communication is defined in such a way that *all* behaviors are to be included in the category, there is nothing to be gained from labeling this all-inclusive category as "communication" rather than simply as "behaviors" in an interpersonal context. Incidentally, it is not always clear whether investigators who hold this definition mean that all behaviors which occur in the context are communications, or that only those behaviors of which they can make sense, which reoccur, and which elicit predictable responses from the other people in the context will be included.

There is a group of investigators who define nonverbal communication as those behaviors which co-occur with the verbalizations in an interpersonal or face-to-face interaction. The issues raised for investigators who define the nonverbal communication instances by the interpersonal context are equally applicable here, even though this group has further limited the set of nonverbal behaviors labeled communication to only those behaviors that accompany the verbal communication. Given that the instances of concern involve explicit communication (i.e., something is being made public via verbal content), the probability is increased that at least some of the movements occurring may also be communications. The problem, however, is that this group of investigators, like the previous groups, gives no independent operations for deciding which if any of the nonverbal behaviors involve code, encoding, and decoding.

The work of Efron (1941) seems to exemplify best the approach of this group. Starting with the observation that the hand and arm movements of two groups from different language backgrounds also differed in terms of size, shape, and tempo, Efron began to investigate the movements themselves. By limiting himself to movement co-occurring with verbalizations, Efron was able to posit the referents of some of the movements, pointing out that these movements could refer to objects, actions, ideas, relations between objects or ideas, or to the flow of ideas. To the extent that Efron did not check his inferences about these movements as code elements, standing for referents, his inferences have the same limitations as those of other investigators who have studied communication solely from an observer's perspective. Nevertheless, to the extent that the behaviors observed were exhibited by several members of a given group, that these behaviors seemed to occur only in the context of a verbal exchange, and that they seemed to be related to the ongoing verbalizations, such behaviors seem more likely to be possible instances of communication from our perspective than the behaviors selected for study by most other investigators. They appear even more promising as communication instances, since several independent observers have noted similar behavior patterns in different groups and at different times (e.g., Ekman & Friesen, 1969; Saitz & Cervenka, 1962; Wiener, 1968). Despite this promise, however, systematic work is still required to make it possible to conclude that this behavior set is a code and part of a communication set. Later in this paper we outline one possible approach toward this end.

In sum, the investigators of socially patterned behaviors have described a number of behaviors and behavior sequences which could well be communications by our definition. They have not, however, clearly discriminated learned behaviors from behaviors which have referents, or habitual response sequences from sequences peculiar to communicative exchanges. They fail to make these discriminations primarily because they, like the investigators of behaviors of the individual, seem to focus on observer's responses to or inferences about the behaviors, rather than considering the subject's use of the behavior to refer to a momentary state or experience. The investigators of socially patterned behaviors have generally justified calling the observed behaviors "communications" by pointing out that the behaviors in question show many of the characteristics of verbal communication: regularity of occurrence, patterns of elements or organization, or response-eliciting characteristics. These investigators have not yet demonstrated that the behaviors in question share with verbal communication the function of referring to something else. The end result is that in a sense these investigators have defined communication in such a way that almost anything could be considered a communication. If communication is to be considered as a subset of behaviors in general and of socially patterned, interactive behaviors in particular, some further definition is required which will make possible the discrimination of behaviors which serve a function similar to that served by verbal communication—namely, the making public of experience via a set of socially shared signals—from those that do not. As we have noted, this discrimination will require some consideration and investigation of encoding, for we hold that the distinction between inference making and decoding cannot be made without some independent evidence of encoding.

An Encoding Approach

In a context which has belabored the point that there have been all too few attempts to investigate communication from an encoding perspective, it is encouraging to note that Ekman and Friesen (1969) seem to be moving toward making distinctions similar to the ones discussed here. Starting with and elaborating somewhat on Mahl's (1968) earlier distinction between informative and communicative behavior, Ekman and Friesen now define communicative nonverbal behavior as

> those acts which are clearly and consciously intended by the sender to transmit a specifiable message to the receiver [in contrast to] informative nonverbal behavior [which is defined as] those acts which have some decoded meaning in that such acts elicit similar interpretations in some set of observers [p. 55].

Within these definitions, there is a clear attempt to separate activities on the part of the behaving subject from activities on the part of the observer; a necessary first step from our perspective in making a sign–communication distinction. Unfortunately, however, while Ekman and Friesen's definition clarifies the sign–communication distinction in some respects, the further explication of these terms introduces ambiguities and confusion in other respects. For example, in these very definitions, Ekman and Friesen do not distinguish the activity of decoding from the activity of inference making, using the terms "decoded meaning" and the making of "interpretations" interchangeably. We note again that any behavior can be a source of information on the basis of which observers may make inferences, but that only some behaviors are decoded. Insofar as Ekman and Friesen do not include a notion of code in their concern about informative events and inference making, the concomitant use of a term such as "decoding," which would seem to imply a code, can only lead to confusion for a reader.

Similarly, in their definition of communication, Ekman and Friesen (1969) seem to imply that encoding on the part of the subject is a prerequisite for considering a behavior to be communicative. We would agree with this prerequisite. However, while the use of the term "encoding" would again seem to imply use of a code, Ekman and Friesen seem to assign a rather different and somewhat ambiguous meaning to this term, as in the following statement: "The act has an encoded meaning in terms of the regularity of its occurrence with those stimulus events which precede, accompany or typically follow it, or with consistently associated ideation [p. 54]." Part of this statement seems to indicate that Ekman and Friesen consider encoding to involve not the use of a code, as the term implies, but regularities in the association of a set of conditions with a particular emitted behavior. If Ekman and Friesen mean by "regularity" that the conditions of occurrence of the behavior can in principle be specified, then all behaviors could be considered regular (and thus, in the particular conceptual framework, encoded). If the term "regularity" involves a notion of frequency of correlation of conditions with occurrence of behaviors (i.e., a notion of predictability), then every simple conditioned response (e.g., eye blinks prior to onset of air puff) must be considered to be encoded. We would prefer to reserve the term "encoding" to refer to regularity in behavior involving the use of a code—the code being specifiable independent of particular instances of occurrence.

Although Ekman and Friesen use the term "coding," their explication of this term appears to preclude the possibility of identifying a code. Ekman and Friesen offer three ways in which a behavior can be related to a referent (i.e., coded), namely, (*a*) arbitrarily, in which the relationship of the behavior to the referent is one of consensus; (*b*) iconically, in which the behavior resembles some attribute or characteristic of the referent; and (*c*) intrinsically, in which the behavior has no referent other than itself. To the extent that Ekman and Friesen seem to imply that coded behavior is used to signify some referent, it is difficult to see why they have included, in coding, behaviors which have no referents other than themselves. Again, in our framework, a code is taken to be a set of behaviors which have referents other than themselves; with this concept, it makes no sense to include, as coding, behaviors which do not have such referents. To imply that every behavior is coded in some way, as Ekman and Friesen seem to do, dilutes the importance of both the concept of code and the concept of coding, and leads once again to the possibility of including any and all behaviors as part of a communication system.

Without an explicit concept of code to aid in the sign–communication distinction, Ekman and Friesen seem to have fallen back on the conscious intentions of the subject as the criterion for considering a behavior to be communicative. Even though Ekman and Friesen recognize that the concept of intention is fraught with empirical problems, they imply there is no apparent alternative.

Since it is not apparent to us how the intention of the communicator can be known by an observer, this concern seems at best irrelevant. If intentionality is defined by asking the communicator to indicate his intentions, then only behaviors which he says he intended as communication can be designated as communicative. In this instance, then, there is no problem. If, however, the observer maintains that the speaker may be unaware of his intentions (i.e., unconscious intention), mistaken about his intentions, or dissimulating, then there is no basis in the behaviors themselves for distinguishing between communicative (intended) and noncommunicative (nonintended) behaviors. In a larger context, it is clear that the distinctions between the truth, a lie, and a mistake are decided on grounds other than the behavioral event or the person's report. To cite a classic example, was Joan of Arc telling the truth about her visions, was she mistaken about what she saw, or was she lying? As is obvious in this instance, the observer must assert his own assumptions in drawing his inferences about the truth value of an event. Until Ekman and Friesen offer some satisfactory epistemological basis for determining the intentionality of the communicator, we see no reasonable solution to the problems inherent in a definition of communication which includes an undefined term, intention. By restricting their use of intention to include only conscious intentions, Ekman and Friesen avoid some of the problems noted above, and state in a footnote that they hope to arrive at some empirical criteria for distinguishing a consciously intended behavior from other behaviors. However, to the best of our knowledge, the specific criteria have not yet been explicated.

Despite ambiguities in the use of the terms coding, encoding, and decoding, and the lack of an explicit concept of code, Ekman and Friesen have begun to specify sets of behaviors which, in their frequency of occurrence, are at least candidates to be considered as part of a communication system. That these latter sets of behaviors have also been delineated by Efron (1941) and have been included in a system of categorization we will offer would seem to indicate that the behaviors may be independent of particular encoders and decoders, thus meeting a second criterion for consideration as part of a communication system.

Incidentally, Ekman and Friesen (1969) have noted (as has Scheflen, 1964a) a class of behaviors which they call regulators. As described by these investigators, this set of behaviors seems to serve the unique function of regulating the communication interaction itself. The set of behaviors which Ekman and Friesen call regulators seems to include the kinds of eye contact behaviors studied by Exline, Gray, and Schuette (1965), what Hall (1963) has called proxemics, or what Scheflen (1964a) has called general interaction patterns. If these behaviors do in fact regulate verbal communication, their occurrence can serve as the empirical basis for holding that encoding and decoding are occurring nonverbally. We hope in a later section of this paper to explicate more fully this possibility. However, in the absence of an explicit specification, by Ekman and Friesen or anyone else, of the set of behaviors which serve to regulate a verbal or nonverbal communication, it is all too easy to include as regulators all behaviors of a person to which another person responds or might respond—as some investigators have done.

An Alternative Approach to Nonverbal Communication

Having considered the problems associated with both decoding approaches to communication and an encoding approach with a focus on intention, let us now propose an alternative conceptual model which we hold will lead to a somewhat different research strategy for studying nonverbal communication than has apparently been used thus far. Like other investigators of communication, we are faced with the problems of (*a*) developing a definition of communication, (*b*) operationalizing the concepts in the definition, (*c*) selecting behaviors to study, and (*d*) explicating a research strategy to explore the implications of the proposed approach.

We have thus far in the paper rather loosely defined communication as the making public of experience via a shared code for encoding and decoding. The major concepts at issue in this definition are code and encoding, for only after these concepts are more fully explicated can the concepts of decoding and communication be understood within our perspective. We start with a definition of code which is essentially similar to a dictionary definition of code as a set of arbitrary components which have referents other than themselves, for example, words, Morse code (see Footnotes 4 and 6). From this definition, it follows that a dictionary of such elements and their referents could in principle be compiled. Such a dictionary would be expected to change over time, however, in the same way that a dictionary of verbal forms changes. Further, while we could reasonably expect some behaviors to have more than one possible referent (as in verbal language some words have more than one referent), we assume that the number of referents for any given behavior would be finite and in principle specifiable, just as the several meanings of a given word are finite and specifiable. We further assume that the specific referent of a particular behavior in a particular context may be determined by its relationship to the rest of the communication matrix, again just as the particular meaning of a word in a given instance can almost always be specified on the basis of its relationship to other code components in the communication matrix. For example, the word "rage" has different meanings in the sentence "The man flew into a rage" and in the sentence "The miniskirt is the rage." The specific meaning of the word rage in these sentences depends on its relation (syntactic-semantic) to other components. These assumptions about symbols and referents are clearly at odds with a view which holds that the "meaning" of any behavior can be known only through knowledge of the total context in which it occurs, or with a notion of idiosyncratic code.

Since we assume that the relationship between the code component and its referent is most often arbitrary, that is, any behavior A, B, or C can be used to refer to any experience X, Y, or Z, as long as the participants agree that a given behavior (e.g., A) stands for or refers to a particular experience (e.g., X), we would not expect any nonverbal communicative behaviors to have universal significance, as some investigators apparently imply (e.g., Ekman & Friesen, 1971). Rather, we would expect cross-cultural differences and even subgroup differences in (*a*) the experience to be made public, (*b*) the nonverbal behaviors used to make experience public, and (*c*) the relationship between experience and the particular nonverbal behaviors. To find invariance of behavioral form and referent across cultures would to us be prima facie evidence that the behavior is not part of a coding system. We could agree that some forms of behavior (e.g., facial variations) might be more likely than others to be used in making public some sets of experience (e.g., mood state). However, a behavior whose significance is found to be invariant across cultures (i.e., observers attribute the same significance to the behavior in all cultures) would be difficult to include as a code component, which for us involves an assumption of an arbitrary relationship between symbol and referent. Rather, we would subsume such behaviors under the rubric of stimulus-specific behavior sequences—even if the behaviors occur under different conditions in different cultures (as suggested by Tomkins, 1962, 1963; and investigated by Ekman & Friesen, 1971; Ekman, Sorenson, & Friesen, 1969; Izard, 1968, 1969).

We assume further that the relationship between behavior A and experience X, for a particular communication group, must at the start of the relationship between A and X (or in developing any formal code) have been made with some awareness, although the experience of self-conscious use of A to stand for X may dissipate for the group as well as for the individual with increasing association between A and X. That is, we view this shift to nonawareness as being no different than the equivalent change in awareness in driving a car. As experienced drivers, we are typically not self-conscious or aware of the movements of hands, arms, and feet which are involved in steering in a particular direction—rather, we focus on the direction to be steered and the appropriate movements occur almost automatically. That these behaviors were once in awareness and learned, however, can be testified to by anyone who remembers learning to drive. We hold that the self-consciousness involved in communicative behaviors is evident only for relatively inexperienced communicators. Anyone who has become proficient in the use of a second language will recall the difficulty he once had in finding the proper word to fit a particular referent—an example of self-consciousness in encoding which diminishes with increasing experience.

As we have pointed out, other investigators have also held either that nonverbal behavior constitutes a code which is decipherable, or that behaviors involved in communication must be socially shared. We would not, then, expect much disagreement with a definition of communication which includes the use of a shared code in which the behaviors constituting the code refer to experiences other than themselves. The difficulty and disagreement, then, if there is any, would be about the operations employed to infer the instance of code and the related concepts of encoding and decoding. In sum, if a set of behaviors will be considered to constitute a code, it must be demonstrated both that the behaviors have referents and that the referents of the behaviors are known and used by a group; that is, used by at least two persons who emit (encode) the behavior to stand for the agreed-on referent and take (decode) the emitted behavior of the other to stand for the agreed-on referent.

To the extent that for us encoding means using a set of code components which make experience public, we must specify the operations or behaviors which we will take as indicators of code usage. Without such specification of behaviors, it is all too easy to assert code usage whenever we see code components. Such an assertion might be acceptable if the code components were independently specified and agreed on. However, since nonverbal code components have not been specified and agreed on, investigation of encoding via nonverbal forms requires some set of operations which make it reasonable to infer code usage.

The method we employ to operationalize code usage seems analogous to the procedures suggested by Hofstadter (1941) under his rubric of "objective teleology."[19] The following extended quotation from Hofstadter seems to explicate best the logic of his approach.

> . . . from the methodological standpoint it seems advisable to develop a way of identifying purposive action without linking it to any specific explanatory hypothesis concerning the internal structure of the purposive agent . . . . As in any problem of identification the best procedure consists in observing and analyzing cases of action which are definitely recognized or labelled as teleological, in order to make explicit the common traits which they exhibit. . . .
>
> There is, then, no initial difficulty in locating objective teleological processes in the rough. The problem is, what common traits do these actions exhibit? In particular, where in these actions do we find objectively purposeful character? And the answer is, we never find an objective purpose by itself, but always in association with a certain "sensitivity to conditions" and a fund of "operative techniques" possessed by the actor. To seek for objective purpose alone, without reference to these two factors, is to embark upon an impossible quest. A purposeful action is directed to its end always in a concrete set of circumstances and along paths of connection between antecedents and consequences. Differences between purposive actions rest not only upon differences of ends, but also upon the range and depth of the circumstances or conditions which enter effectively as well as upon the scope of the connections of antecedents and consequences actually operative. This becomes especially clear when considered from the viewpoint of prediction. In the case in which a teleological action is not as yet complete, if I merely assume that the actor is pursuing a cer-

---

[19] Though our method is similar to Hofstadter's, our concerns are not with the philosophical problem of teleology. We are indebted to Leonard Cirillo for pointing out to us the similarity between our approach and Hofstadter's.

> tain end I can not from this predict the methods and procedures it will employ nor the occasions and matters upon which it will employ them. I need to assume that it is "sensitive" to certain circumstances in specific ways; and I need also to assume that it has a fund of operative techniques, of abilities to connect antecedents and consequences, i.e., that it will respond to a specific condition to which it is sensitive in a specific way. It is possible that many different actors should have similar ends and yet behave in radically diverse ways. The differences of behavior are not predictable, and do not follow, from the similarity of ends, but only from the ends plus the differences in sensitivity to conditions and in the funds of operative techniques available to the actors. But since, in predicting, I need to be able to predict differentially, it follows that my assumption of end alone does not provide me with an adequate empirical hypothesis. On the other hand, if to the assumed end I enjoin a specific sensitivity to conditions and a fund of operative techniques, I do get an empirical hypothesis capable of differential prediction and therefore susceptible of confirmation or disconfirmation. . . . To discover operative techniques in teleological process in general requires a number of contexts of activity in which the subject of investigation appears and in which ends and sensitivities to conditions vary. Regularities in the choice of similar means for similar ends under similar conditions, different means for similar ends or similar means for different ends under different conditions, different means for different ends under similar conditions—all these indicate technological principles in operation, operative techniques. . . .
>
> The unitary attribute of the teleological actor is not the possession of end alone, or sensitivity alone, or technique alone, but all three in inseparable combination. However, it is also true that although they can not be separated in the unitary attribute they may nevertheless be analyzed out independently by the use of a plurality of acts of the same agent. It is possible to discover the fund of generally operative sensitivity and technique possessed by the subject of investigation by observing its behavior in varying contexts, where there are convincing reasons to believe that ends vary, and segregating what is constant in sensitivity or in technique. By analogous methods one can study the general character of its ends by varying the contexts so as to call for variations in sensitivity and technique. It is thus possible to discover constancies in any one of the three factors with variations in the others.
>
> It must be emphasized that ascription of a unitary objective teleological attribute, consisting of end, sensitivity, and technique, to an object is in no way an explanation of its behavior. It is simply a shorthand way of describing that behavior. The behavior exhibits a pattern of directedness, passing through certain specific circumstances within a context, and proceeding by certain connecting paths. The directedness is a directedness of the paths through the contextual circumstances. Directedness without paths and circumstances is nothing. Paths in no direction and connecting no circumstances are nothing. Circumstances not located on paths leading in a direction are nothing. "End" is a name for the directedness, "sensitivity" for the circumstances, "technique" for the connecting paths, when taken in relation to the agent. The problem of explaining the action arises only when one asks whether a correlation can be made between these elements of the actions and structural-dynamic properties of the actor. The existence of the teleological process is sufficiently attested when a hypothesis about end, sensitivity, and technique is confirmed [pp. 32–35].[20]

Hofstadter appears to be arguing that an event can be considered goal directed (*a*) if some observable set of criteria can be stated (in this instance the sensitivity of the doer, the conditions in which the act occurs, the means, and some end state), and (*b*) if modifying any one of these leads to a predicted change in the others. The underlying principle involved in his approach to investigating behaviors or events which have been traditionally defined in subjective terms, then, seems to be that of deriving a set of operational events which are posited as being related, and then checking outcome predictions made on the basis of this posited interrelationship when one of the relationships among the designated events is changed. If the predictions are confirmed in a variety of instances in which the relationships among criterial events are manipulated, then Hofstadter holds that goal-directed behaviors have been investigated both empirically and objectively.

As we have noted above, our method for operationalizing encoding or code usage involves a strategy similar to that proposed by Hofstadter. That is, code usage will be posited on the basis of a set of interrelated predicted findings when manipulation of the experiences to be made public, of the forms available to the addressor, or of the responsiveness of the addressor to communication conditions results in the predicted effect on the emitted behaviors as a function of such manipulations.[21] Specifically, this method involves five interrelated steps. First, we start with behaviors which are consensually taken to be an instance of code usage—namely, verbal language behaviors.[22] Second, we identify a set of behaviors associated with verbal language behaviors. The occurrence of this set of behaviors is taken to be independent evidence for, and consistent with, the assumption that encoding is taking place in verbal behaviors. That is, we will identify behaviors from which we can conclude that the encoder's verbal behavior is systematically modulated as there is a change in the experience to be made public, in the conditions of the communication, or in the response patterns of the decoder. Third, we demonstrate that in the same way that predictable introductions take place within verbal language (i.e., variations in words occur as a function of systematic manipulation of the experience to be made public, of the conditions of the communication, or of the decoder's response patterns), predictable nonverbal behaviors will be introduced when the verbal possibilities are constrained. Fourth, we show that the introduction of these nonverbal behaviors does not result in changes in those decoder behaviors which we take to be indicators of decoding or understanding in verbal exchanges. Fifth, we argue that if these nonverbal behaviors are introduced predictably, and if their introduction does not result in a significant change in decoding indicators, then the nonverbal forms introduced can be considered to be substituted for verbal forms, and thus can themselves be considered as components of a nonverbal code whose emission involves code usage.[23]

---

[20] Reprinted with permission from an article by Albert Hofstadter published in the *Journal of Philosophy*, 1941, Vol. 37. Copyrighted by the *Journal of Philosophy*, 1941.

Four classes of behavioral events can be identified whose occurrence coincides with our assumption that verbal language involves code and code usage. Two of these classes, search and correction, appear to be indicators of a code, that is, a standard against which emitters check their verbalizations. The other two classes, consistent with an assumption of code usage, are (*a*) regulators, one set emitted by the addressee to indicate decoding and another set emitted by the addressor to check on whether decoding is occurring, and (*b*) message modulations, which are responses to variations in the event to be made public, to the conditions of communication, and to the regulators or other behaviors of the addressee. These four classes of behaviors appear in both verbal and nonverbal forms during verbal interactions. In describing these classes of events, the four classes will be presented as if they were independent occurrences, but for the most part two or more usually seem to co-occur, either sequentially or concurrently.

Search typically involves a pause greater than the stops required for paralinguistic designation of word, phrase, or sentence; such a pause may also occur at points other than paralinguistic junctures in the flow of verbal communications. The pause is usually accompanied by the speaker looking away (usually up) while maintaining a constant tonal pitch (the pitch neither drops, as at the end of a declarative statement, nor rises, as at the end of a question). At the end of the pause, the speaker reinstitutes the flow of verbal communication. Verbal forms of search may include such statements in the pause as, "I can't think of the right word," or "What's that word that means the same as—?" We have not noted a significant number of instances of search for nonverbal components, except under special circumstances such as charades or sign language.

---

[21] As we use the terms code usage and encoding, they do not imply goal-directed behaviors, nor do they imply an objective or subjective teleological explanation. At this moment we hold only that if verbal language exchanges are taken to be an exemplar of code usage and encoding behavior, then we can specify a set of operations which will make it possible to hold that particular nonverbal behaviors are code components and part of a communication system in the same way words are. This approach we propose is also analogous to that offered by Cronbach and Meehl (1955) in their discussion of construct validity.

[22] As we have noted earlier (see Footnotes 4 and 6), for us language includes a code, rules for the emission of code components, and rules governing organization of code components.

[23] We recognize the similarity of our substitution criterion to Mahl's (1968) definition of communication. The difference between the two approaches is a function of how substitutability is determined from a decoder versus an encoder perspective.

Correction is usually noted when a speaker changes an element or component in his ongoing flow of speech. It is most evident in a verbal exchange when an emitter says, "I meant to say ambiguous, not ambivalent," or when an addressee says something like "You mean ravenous, not ravishing, don't you?" When the emitter makes a correction, it is often preceded by search; correction by the addressee typically occurs immediately or as soon as the speaker pauses (i.e., at the end of the thought). Examples of addressee correction of nonverbal forms in communication sometimes occur in response to tonal variations, as when a mother says to her child, "Don't talk to me in that tone of voice," or by an emitter when he says "I didn't mean to say it that way," and repeats the message, changing the tonal variations. As noted above for search, correction of other nonverbal components (e.g., gesture) is also rare, again except in special situations such as charades or sign language. It may well be that the infrequent occurrences of search and correction for nonverbal forms is a function of the relatively small number of components in these nonverbal codes.

Regulators are behaviors which we hold are taken by addressees and emitters as signs that encoding or decoding are occurring, although we do not include these behaviors in the code itself. The set of behaviors we call regulators is more restricted than the set implied by some investigators (e.g., Scheflen, 1964a), in that our set includes only those behaviors which are emitted by most encoders and decoders of a communication group and in most, if not all, communication situations. The set of regulators which we describe here derives in part from the thought-provoking and intriguing study of particular regulators by other investigators (e.g., Exline et al., 1965; Hall, 1963). Although we try to describe each regulator independently, these behaviors also seem to occur in varying combinations. We have no reason to believe that the behaviors we describe are the only regulators which occur, and the set below is not considered to exhaust the category.

Addressor regulators are behaviors emitted by the speaker which appear (*a*) to serve as cues for the addressee that encoding is occurring and continuing, (*b*) to check on whether listening and decoding are occurring, and (*c*) to indicate when the addressee is to speak. The following behaviors have been identified:

A. Eye movement. The addressor typically looks at the addressee at the end of completed thoughts during the flow of speech, with or without a short pause. If the addressee makes eye contact and/or nods at those times, the emitter seems to take this as evidence that the addressee is listening and decoding, and the emitter continues speaking. If the addressee makes eye contact, but with a blank or quizzical look, the speaker typically rephrases or elaborates his message. If the addressee does not make eye contact at these times, the speaker either repeats his message with increased volume or introduces other cues (e.g., throat clearing) which evoke listening on the part of the addressee, or the emitter ceases speaking altogether. However, if the emitter pauses for longer than usual without initiating eye contact, but moves his eyes to the side or up instead, it seems to be taken by the addressee as an indication that the addressor is thinking or searching and will continue his speaking following the pause.

B. Vocal variations. Two particular kinds of vocal variations, inflection and pause, can be considered as regulators. While some of these variations also have paralinguistic significance, we discuss them only as regulators. An upward tonal inflection (as in a question) or a downward tonal inflection (as in a declarative statement), when accompanied by an extended pause, seems to be taken as requiring some response from the addressee—either speaking or some indication of decoding (see below). However, a pause and extended eye movement to the side or eyes rolled up, even with inflection changes, seem to be taken by the addressee as an indication that the emitter is thinking or searching and will continue to speak; no response is required by the addressee.

Addressee regulators are behaviors emitted by the addressee which seem to be taken by the addressor as cues that decoding, understanding, agreement, or disagreement are occurring.

- A. Eye contact. The addressee is expected to maintain relatively constant eye focus on the addressor while the latter is speaking. When the addressor pauses and checks the addressee for eye contact, at least four possible responses by the addressee can occur.
    1. The addressee moves his eyes upward without speaking; this behavior seems to be taken as a cue that the addressee is thinking and will respond subsequently. The original emitter waits.
    2. The addressee maintains a quizzical or blank look without speaking, usually taken as a sign that the addressee has not understood the message, and responded to by the emitter with a message variation.
    3. The addressee maintains eye contact with a slow rhythmic head nod. This behavior seems to be taken by the addressor as evidence of decoding or understanding, and the addressor can continue speaking. This nod may be contrasted with a sharp staccato nod, which seems to indicate agreement in addition to understanding.
    4. The addressee maintains eye contact and smiles. This behavior seems to be taken as a cue that the addressee understands but has nothing to add.
- B. Other addressee regulators.
    1. A raised eyebrow or a frown seem to be taken as indicating disagreement or questions; the addressor typically reiterates, corrects, or elaborates his message.
    2. A sudden intake of breath by the addressee, co-occurring with the opening of the mouth as if to speak, and sometimes accompanied by a postural shift forward, seems to indicate that the addressee has something to say. This sequence is followed by an abrupt end to the addressor's speech or, if speech continues, by an increase in pitch and/or volume and/or rate of speech. This entire latter sequence is sometimes accompanied by an upheld palm as a pantomime of "stop," and all of this seems to indicate that the emitter has responded to the interruption cue, but will override it and make his point.

These regulators have been observed in a wide variety of situations and for many different addressors and addressees. They are a most interesting phenomenon in communication and deserve more systematic study in their own right. For our purposes here, we only propose to demonstrate that among other manipulations, the control of regulators will lead to predictable variations in nonverbal behavior as well as in verbal communication.

Message modulation within the context of verbal code usage is taken to mean predictable variations in form or style of utterance, or in the components used. These verbal variations are consistent with the assumption of code usage in that the modifications seem to provide evidence that the speaker is taking into account, in his verbalization, changes in the experience to be made public, the conditions of the communication, and the characteristics and responses of the addressee. For example, a young male in describing his activities with a date on the previous evening will use different components when relating the events in a locker room and when speaking to the same addressee in a restaurant. As another example, upon hearing an addressee's foreign accent, or when speaking to a child, an emitter will speak more slowly and clearly and may also use words that are more common than the words he would use with a native adult addressee. As still a further example, a man who complains to his fellow workers using one set of statements will reword the complaint when speaking to the president of his company.

We have briefly outlined a set of behaviors associated with spoken verbal language which will be taken as independent evidence of code usage. Our approach now requires us to define communication more explicitly and then to (*a*) specify a set of behaviors which are hypothesized to be nonverbal code components; (*b*) demonstrate that search, correction, and regulators co-occur with these behaviors; and (*c*) show that the hypothesized code components will occur predictably; that is, they can be seen as substitutions for verbal code components, given the appropriate conditions for message modulations accompanied by restrictions in the use of the verbal code.

In this context, then, communication is operationally defined as the occurrence of encoding and decoding behaviors, particularly as evidenced by regulators. Whether or not understanding does in fact occur, that is, whether emitted behaviors by the communicators have shared referents, may only be known subsequently (e.g., by noting whether subsequent behaviors by both communicators show a correspondence of referents), or by the use of some other criteria (e.g., consensus judgments of other members of the communication group). When mismatches of referents are found, we call such instances code-usage errors. We hold that correction of such errors or mismatches is part of the ongoing learning and elaboration of the code which is available to all members of a communication group. In sum, communication is defined by the ongoing, concurrent behavior sequences of encoding and decoding. Errors in the use of referents are known by some other set of criteria—correspondence of other behaviors or consensus of members of the communication group (e.g., Hymes, 1964).

Given our strategy of prediction, our third step is to denote a set of behaviors—in this paper we outline hand and arm movements—which are possible code components. There are several criteria that such behaviors must meet if we are to consider them as possible components in a communication code, as opposed to reactions to some immediate stimulus or as opposed to a socially patterned behavior (i.e., patterned behavior learned in the process of acculturation):

A. The behaviors must be emitted by the particular communication group studied. This criterion eliminates idiosyncratic behaviors and meets the requirement of socially shared behaviors.

B. The behaviors must occur in several different contexts. This criterion is likely to eliminate some behaviors which are reactions to a specific set of stimulus conditions.

C. The behaviors must be more likely to occur in verbal contexts than in any or all other contexts. If a behavior (e.g., scratching) can occur in any context—that is, with or without an addressee—it is difficult to accept it as a possible component in a communication code.

D. The behaviors as code components should encompass a relatively short time duration. This criterion serves to focus on ongoing experience rather than on socially prescribed patterns of behavior or on behavior related to personality styles (e.g., the wearing of rings or the handling of teacups).

These four criteria for selecting behaviors to be studied make it possible to eliminate a great number of other behaviors which have been considered in the non-verbal-communication literature. For example, these criteria eliminate from our concern the autistic behaviors studied by Krout (1935a, 1935b) and by Mahl (1968), many of the situation-specific behaviors reported by Scheflen (1965) and by Birdwhistell (1970), and many of the habitual, learned social behaviors presented by Goffman (1959).

Further we have found it helpful to utilize the concept of "channel" in selecting behaviors for study. Channel has been defined (Wiener & Mehrabian, 1968) as any set of behaviors in a communication that has been systematically denoted by an observer, that is considered by that observer to have coding possibilities, and that can be studied (at least in principle) independently of any other co-occurring behaviors. The designation of a set of behaviors as a channel in communication is quite arbitrary and is based only on some division of the communicative events which has been found useful for investigation. In essence, then, the designation of a channel has often been dictated by pragmatic considerations or conventionally accepted distinctions, as when body movements are set apart from vocalizations as constituting different classes of events. There is no implication that these specifications of the behaviors as channels preclude other ways of subdividing the communicative event, as long as the perspective of code usage is maintained. Further, no attempt is being made to reify the concept of channel. In sum, then, a channel is a set of behaviors arbitrarily delineated either on the basis of conventional distinctions or on the basis of some other organizing principle.

Using the concept of channel and the criteria for selecting behaviors to be studied, we have isolated the following movements of the hands and arms (gestures). The gestures, along with their posited significance and functions, are listed below. Again, it should be noted that the categories and/or the particular gestures have been described by one or another of the previous investigators in this field (e.g., Efron 1941; Ekman & Friesen, 1969; Krout, 1935a, 1935b; Saitz & Cervenka, 1962).

A. Pantomimic gestures.

1. *Formal pantomimic gestures* are stylized movements of the arms and hands for which there is a culturally prescribed consensual meaning. Typically, such gestures have an object or event significance and function as nouns do in the verbal channel. Formal pantomimic gestures are almost invariant in meaning despite variations in the context, situation, time, or addressee; this relative invariance in meaning is similar to the relative invariance in meaning of words in the verbal channel. Formal pantomimic gestures can take the place of words; when they co-occur with words, the function of the gesture is that of emphasizing by adding redundancy to the verbal communication. Examples of formal pantomimic gestures are waving goodby, depicting two wavy lines in the air to designate a well-proportioned female, "the finger," and forming a circle with the index finger and thumb, remaining fingers extended to signify "okay" or "perfect."

   The number of gestures which have a culturally prescribed meaning is typically limited within the general population, although some subgroups within a culture may make more extensive use of gestural substitutes for words than other subgroups, and may even develop their own formal pantomimic gestures. For example, in nineteenth-century Russia, the sophisticated balletomane knew that a gesture in which the thumb touched the left, middle, and right side of the forehead meant "king," while a gesture in which the hand encircled the face meant "beautiful girl." Revivals of nineteenth-century ballets still include long portions of pantomime, in which stylized gestures are used to tell the story of the ballet. Other examples are evident in special groups such as an army unit, or pilots, or other groups who communicate over long distances. In games such as charades, gestures which are consensually defined would be classified as formal pantomimic gestures, while other gestures used in charades, which are not consensually defined, fall into the category of improvisational pantomimic gestures.

2. *Improvisational pantomimic gestures.*[24] Although we have no reason to believe that these forms have as much prescribed meaning as do formal pantomimic gestures, meaning seems to accrue in a particular context as a function of *some perceptual attribute* of an object, of an action performed on or with an object, or of the reaction to an object or event. The improvisational pantomimic gestures also have a noun or event quality, but since the meaning ascribed to the gesture is dependent on the situation, context, or content, these particular gestures occur most frequently in a context which has co-occurring verbal communication. For the most part, improvisational pantomimic gestures seem to serve the function of emphasizing, concretizing, or focusing on a particular aspect of the communication occurring verbally. Examples of gestures within this category of communicative movements are drawing a number of circles in the air around one point to depict the concept concentric and swinging an imaginary bat. There is a special set of circumstances in which improvisational pantomimes seem most prominent; that is, when more typical communicative patterns are blocked—for example, when speaking to a foreigner.

B. Semantic modifying and relational gestures. These gestures usually accompany speech and are hypothesized to (*a*) serve the function of modification and specification of the communication, the same function served by adjectives and adverbs in the verbal channel; (*b*) specify the speaker's relationship to the addressee and to his communication; or (*c*) specify the relationship of one aspect of the communication to another aspect of the communication. The occurrence of these nonpantomimic gestures seems to be dependent on the communicative content co-occurring in other channels, but the forms of the gestures, when they do occur, seem to be relatively invariant and appear to have independent meaning. In general, the dimensions of these nonpantomimic gestures, along which variations seem to indicate differences in communicative significance, are (*a*) pointing and its direction, (*b*) orientation of palms, (*c*) semantic forms, and (*d*) location and size of gestures.

---

[24] There is an apparent contradiction in including a set of gestures labeled improvisational under a rubric of shared code elements. Our decision to include these movements as possible communication components derives from an assumption that the particular addressor and addressee share a knowledge of the particular object or event and their attributes; objects and events referred to explicitly in the verbal content. Even though the particular movements themselves are not likely to occur again in other contexts or with other addressees who do not have the shared experience of the object or event, in the instances where they do occur, the referent (the object) is posited to be shared knowledge.

1. *Deictic movements* are pointing movements which can be done with the index finger, two or more fingers, or all fingers. Deictic movements are hypothesized to serve the function of specifying or reducing the ambiguity of the verbal referent. For example, the pronoun "you" can mean a specific person or a generalized other. When accompanied by a pointing to the addressee the meaning of "you" is the addressee; if pointed somewhere else than at the addressee it means "other than you, but not present" or "not I." Pointing gestures are also used to specify the relative positions of objects in space or events in time—"not here, not now" instances.
2. *Orientation of palms*—The position of the palms in a gesture indicates the speaker's relationship to the addressee—for example, mutual versus nonmutual—and/or about the speaker's attitude toward his communication—certain versus uncertain. Palms up is equivalent to uncertainty or to "I think" or "I believe" or "It seems to me" in a verbal statement, and adds for the addressee the message that the issue need not be pursued since noncertainty is indicated. Palms down indicates an assertion with the speaker again communicating that the subject matter is not open to question—equivalent verbal statements are "clearly," "absolutely," "without doubt," etc. Palms out (i.e., facing toward the addressee) is a statement of assertion and is equivalent to the statement, "I shall say it" or "Don't interrupt." These last gestures also seem to function as regulators of the communication interaction.
3. *Semantic forms*—Five gestures or gestural patterns have been isolated: (*a*) circling gestures, (*b*) oscillation gestures, (*c*) arhythmic "chopping" gestures, (*d*) rhythmic "chopping" gestures, and (*e*) expansion-contraction gestures. Each of these will be described briefly and its hypothesized, communicative significance posited.

   a. Circling gestures—a slow, continuous series of circular movements of any part of the hand including the arm which is hypothesized to indicate nonspecificity, globality, or generality of the verbal components in the communication. In words, this gesture is posited as indicating "I mean more than the specific words I have used," or "The specific referent term is only one general attribute of the event I'm describing."

   b. Oscillation gestures—a series of slow to moderate, semicircular, back and forth movements of the hand which seem to indicate "either or" or "one or the other" of two components, or "on one hand versus on the other hand."

   c. Arhythmic chopping gestures—linear, staccatolike movements of the arm and hand, usually performed in a limited series of one or two in which the same plane is maintained throughout the movement. The chopping gestures seem to be associated with the speaker's experience of discreteness of the elements or contents of his verbal communication. For example, a verbal statement in the form of "We have to consider A and B and C" is likely to be accompanied by a series of three "chops"—one for each component. The verbal statement in the form, "We have to consider John and his behavior" with a single chop at the end is interpreted to mean only his behavior—discreteness indicator—versus two chops where John in general is to be considered and his behavior is to be considered separately.

   d. Rhythmic chopping gestures—also linear, staccato movements, usually *more* than one or two, which have a rhythm independent of the rhythm of the co-occurring verbal content. In this form, the staccato chop is interpreted to mean emphasis rather than discreteness.

   e. Expansion-contraction gestures—a slow expansion and contraction of the arm in which the addressor looks as if he were playing a concertina. Typically these movements indicate a size or extent dimension of the referent carried in the verbal channel.

   In sum, the semantic forms seem to serve the function of specification and modification served by adjectives and adverbs in the verbal channel of communication; the palms seem to indicate the speaker's relationship to his communication, the addressee, or the relationship of one part of his communication to another part of his communication. However, much of the information communicated through the gestural channel appears to be redundant with that communicated through other channels. For example, tonal changes are effective indicators of the emphasis a speaker wishes to place on any portion of his verbal communication; facial expressions can communicate doubt or uncertainty as easily as gestures. The chief communicative functions of the gestural channel appear to be to specify the referent of an ambiguous verbal statement, to specify the addressor's relationship to his verbal communication, and to indicate intensity or emphasis by introducing redundancy into the message.

It is further hypothesized that the size and location of the gesture are co-varying attributes which indicate the speaker's experience of the centrality or the importance of his gestural communication. A small gesture made in the plane directly in front of the speaker is taken as equivalent in communicative significance to a larger gesture made more peripherally, the view being that a gesture must be seen to serve a communicative function. In line with this hypothesis, it would seem that the speaker's experience of the centrality or importance of the communicative component carried gesturally will determine both the size and location of the gesture. From this latter hypothesis comes the derived statement that variations in the size and location of the gesture are associated with variations in the emphasis or intensity placed on the gestural component of the message.

Preliminary research has been concluded using this set of descriptive categories. In one study, untrained subjects were shown videotaped exemplars of each of the gestures and were asked what they took each gesture to mean; the subjects, as predicted, assigned to the gestures the appropriate meaning category (Rubinow, 1970). Interestingly, there were some intriguing individual differences in the reliability of these categorizations. In a second study using videotape recordings of interviews without the sound on, trained judges were able to discriminate the gestures from one another with a high degree of reliability (Novick, 1967).

Up to this point we have (*a*) specified a set of events (search, correction, regulators, and message modulations) whose occurrence can be taken as independent evidence of code and code usage, (*b*) elaborated a set of criteria for selecting behaviors as possible components in a nonverbal communication code, and (*c*) described some hand and arm movements as exemplars of the kinds of behaviors which might be investigated in nonverbal encoding. What remains to be done, then, is to outline a possible research strategy for determining whether the posited nonverbal code components occur predictably as a function of changes in the experiences to be made public, of the conditions of the communication exchange, of the addressee regulators, or of the subsequent verbal behaviors of the addressee when the addressor's use of verbal forms is restricted. In addition to demonstrating the predictability of the occurrence of the nonverbal forms, it is necessary to demonstrate that the addressee's decoding response to the occurrence of the nonverbal forms is equivalent to his decoding response to a verbal utterance where use of the verbal forms is not restricted.

Below is an outline of some possible manipulations which could be attempted within our proposed strategy.

A. Manipulation of the experience to be made public.
  1. A subject is required to talk about a subject matter he knows well and about a subject matter he knows less well (e.g., a psychologist is required to talk about some subject in psychology and about some subject in physics). We would predict an increase in the number of uncertainty gestures (i.e., palms up) and of vagueness or generality gestures (i.e., circling) when the subject talks about less familiar subject matter.
  2. A subject is required to say statements with which he agrees and statements with which he disagrees. We would expect more rhythmic chops and palms down in the former case than in the latter case. Incidentally, we would also expect that the gestures which co-occur with agreement statements would be larger and/or would be located in a more central plane.

B. Manipulation of the conditions of communication.
  1. The number of channels available by which a subject can communicate could be varied, as when, for example, a subject is required to speak in a foreign language he has just begun to learn. We assume far less verbal fluency for the second language than for a first language and would predict increased numbers of pantomimes, palms up, and circling gestures for the second language condition.
  2. A subject is to communicate face-to-face and via telephone to the same addressee. We predict more semantic modifiers and pantomimes in the face-to-face condition than for the telephone condition, with a concomitant decrease in verbal explicitness for the face-to-face exchange.

C. Manipulation of addressee variables.
  1. A subject would speak about the same things with two addressees, one of whom did not emit the customary decoding regulators. We predict that in the absence of

these regulators, the addressor will begin to use more gestures, particularly more pantomimes and more palms up, and will change the verbal components radically.

2. A subject is required to speak to an adult who is identified as a native speaker and an adult who is identified as a foreigner. We predict a marked increase in pantomimic gestures for an addressee who is assumed to have difficulty in decoding. We would predict similar changes for addressees who are children.

The number of possible experimental manipulations and testable hypotheses is dependent only on the investigator's ingenuity. Further, as we have noted earlier, we hold that this strategy should be applicable to the study of any channel in communication, with the stipulations that addressors and addressees be drawn from the same communication group and that each addressor be assigned to two conditions (e.g., restriction on verbal code use vs. no restriction). The responses of the addressees for the two conditions can then be compared for decoding equivalence.

Once code and code usage can be specified for a given channel, it should be possible to investigate (*a*) developmental differences in code usage (e.g., deictive movements and informal pantomimic gestures are expected to occur earlier than the semantic modifying and relational gestures); (*b*) individual differences in code usage (e.g., verbal articulateness and amount of gesture are expected to be inversely related); (*c*) subgroup differences in code usage and responses, as well as in the use of and response to regulators (e.g., Brooks, Brandt, & Wiener, 1969, have found such subgroup differences in response to tonal inflection); and (*d*) differences in code usage for groups with differential psychiatric diagnoses (e.g., obsessive-compulsives are expected to use more gestures of enumeration and relationships; schizophrenics are expected, as nonconsensual communicators, to use few, if any, gestures and to respond to—that is, decode—few gestures).

In sum, our proposed approach attempts to emphasize encoding in the study of communication rather than relying on a decoding perspective, and attempts to operationalize the concepts of encoding, decoding, code, and communication. This attempt will, we hope, make possible more systematic study of nonverbal communication than has generally occurred. It is also hoped that the making explicit of one alternative approach will stimulate others to offer other alternative approaches to the study of nonverbal communication.

REFERENCES

ALLPORT, G., & VERNON, P. *Studies in expressive movements.* New York: Macmillan, 1933.

ALSTON, W. P. *Philosophy of language.* Englewood Cliffs, N. J.: Prentice-Hall, 1964.

BIRDWHISTELL, R. L. *Kinesics and context: Essays on body motion communication.* Philadelphia: University of Pennsylvania Press, 1970.

BROOKS, R., BRANDT, L., & WIENER, M. Differential response to two communication channels: Socioeconomic class differences in response to verbal reinforcers communicated with and without tonal inflection. *Child Development*, 1969, **40**, 453–470.

CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychological Bulletin*, 1955, **52**, 281–302.

DEUTSCH, F. Correlations of verbal and nonverbal communication in interviews elicited by the associative anamnesis. *Psychosomatic Medicine*, 1959, **21**, 123–130.

DITTMANN, A. T. The relationship between body movements and moods in interviews. *Journal of Consulting Psychology*, 1962, **26**, 480.

DUNCAN, S. Nonverbal communication. *Psychological Bulletin*, 1969, **72**, 118–137.

EFRON, D. *Gesture and environment.* New York: King's Crown, 1941.

EKMAN, P. Body position, facial expression, and verbal behavior during interviews. *Journal of Abnormal and Social Psychology*, 1964, **68**, 295–301.

EKMAN, P. Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology*, 1965, **2**, 725–735. (a)

EKMAN, P. Communication through nonverbal behavior: A source of information about an interpersonal relationship. In S. S. Tomkins & C. E. Izard (Eds.), *Affect, cognition, and personality.* New York: Springer Press, 1965. (b)

EKMAN, P., & FRIESEN, W. V. Nonverbal behavior in psychotherapy research. In J. Schlien (Ed.), *Research in psychotherapy.* Vol. 3. Washington, D. C.: American Psychological Association, 1968.

EKMAN, P., & FRIESEN, W. V. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1969, **1**, 49-97.

EKMAN, P., & FRIESEN, W. V. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 1971, **17**, 124–129.

EKMAN, P., SORENSON, E. R., & FRIESEN, W. V. Pan-cultural elements in facial displays of emotion. *Science*, 1969, **164**, 86–88.

EXLINE, R. V., GRAY, D., & SCHUETTE, D. Visual behavior in a dyad as affected by interview content and sex of respondent. *Journal of Personality and Social Psychology*, 1965, **1**, 201.

FELDMAN, S. *Mannerisms of speech and gestures in everyday life.* New York: International Universities Press, 1959.

GOFFMAN, E. *The presentation of self in everyday life.* New York: Doubleday, 1959.

HAGGARD, E. A., & ISAACS, N. S. Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. In L. A. Gottschalk & A. A. Auerbach (Eds.), *Methods of research in psychotherapy.* New York: Appleton-Century-Crofts, 1966.

HALL, E. T. A system for the notation of proxemic behavior. *American Anthropologist*, 1963, **63**, 1003–1026.

HOFSTADTER, A. Objective teleology. *Journal of Philosophy*, 1941, **37**, 29–39.

HYMES, D. *Language in culture and society: A reader in linguistics and anthropology.* New York: Harper & Row, 1964.

IZARD, C. E. Cross-cultural research findings on development in recognition of facial behavior. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1968, **3**, 727. (Summary)

IZARD, C. E. The emotions and emotion constructs in personality and culture research. In R. B. Cattell (Ed.), *Handbook of modern personality theory.* Chicago: Aldine Press, 1969.

KROUT, M. H. Autistic gestures: An experimental study in symbolic movement. *Psychological Monographs*, 1935, **46**(4, Whole No. 208). (a)

KROUT, M. H. Social and psychological significance of gestures. *Journal of Genetic Psychology*, 1935, **47**, 385–412. (b)

KROUT, M. H. Understanding human gestures. *Scientific Monthly*, 1939, **49**, 167–172.

MAHL, G. F. Gestures and body movements in interviews. In J. Schlien (Ed.), *Research in psychotherapy.* Vol. 3. Washington, D. C.: American Psychological Association, 1968.

NOVICK, S. A categorization and comparison of communication movements utilized by two personality types. Unpublished senior honor's thesis, Clark University, 1967.

REIK, T. Men and women speak different languages. *Psychoanalysis*, 1954, **2**, 13–15.

ROBINSON, G. Identification and its relation to empathy, communication, and speech: II. *Journal of Asthma Research*, 1965, **2**, 293–407.

RUBINOW, S. Hand and arm gestures as communications. Unpublished master's thesis, Clark University, 1970.

SAINSBURY, P. Gestural movement during psychiatric interview. *Psychosomatic Medicine*, 1955, **17**, 458.

SAITZ, R. L., & CERVENKA, E. J. *Columbian and North American gestures: An experimental study.* (Carrera 7 23–49) Bogota, Columbia: Centro Colombo Americano, 1962.

SCHEFLEN, A. E. Communication and regulation in psychotherapy. *Psychiatry*, 1964, **27**, 126–136. (a)

SCHEFLEN, A. E. The significance of posture in communication systems. *Psychiatry*, 1964, **27**, 315–331. (b)

SCHEFLEN, A. E. Quasi-courtship behavior in psychotherapy. *Psychiatry*, 1965, **28**, 245–257.

SCHEFLEN, A. E. On the structuring of human communication. *American Behavioral Scientist*, 1967, **10**, 8–12.

TOMKINS, S. S. *Affect, imagery, consciousness,* Vol. 1. *The positive affects.* New York: Springer, 1962.

TOMKINS, S. S. *Affect, imagery, consciousness,* Vol. 2. *The negative affects.* New York: Springer, 1963.

WATZLAWICK, P., BEAVIN, J. H., & JACKSON, D. D. *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes.* New York: Norton, 1967.

WIENER, M. *Gesture in communication.* (USPHS-NIMH Res. Grant 1 RO3 16231-01) Washington, D. C.: United States Public Health Service-National Institute of Mental Health, 1968.

WIENER, M., & MEHRABIAN, A. *Language within language: Immediacy, a channel in verbal communication.* New York: Appleton-Century-Crofts, 1968.

WOLFF, W. Involuntary self-expression in gait and other movements: An experimental study. *Character and Personality*, 1935, **3**, 327–344.

(Received August 23, 1971; revision received December 21, 1971)

Psychological Review
1972, Vol. 79, No. 3, 215–236

ONTOGENY OF MEMORY[1]

BYRON A. CAMPBELL[2]
Princeton University

NORMAN E. SPEAR
Rutgers University

Over the past half century a substantial number of studies have demonstrated that capacity for long-term memory increases during development. This paper presents a brief review of those findings and then examines a variety of mechanisms, both behavioral and neurological, which might account for the increase in memory during ontogenesis.

This paper is addressed to one of the major paradoxes in psychology today. On the one hand, there is a tremendous wealth of literature documenting both the permanence and pervasiveness of experiences occurring during the early stages of development in both animals and man. Adult social behaviors (Harlow & Harlow, 1962), resistance to stressor agents (Levine & Mullins, 1968), birdsongs (Hinde, 1969), sexual behavior (Harlow, 1962), and many other adult behaviors have all been shown to be critically dependent on events occurring early in life. Yet on the other hand, it is apparent from our own experience, from extensive research of childhood memories in adults (Dudycha & Dudycha, 1941; Schachtel, 1947), and from recent research on animals (Campbell, 1967) that memories of specific events occurring early in life are typically more fragile than memories of similar events experienced in adulthood.

Because of the strong implications for human personality, a great deal more attention has been paid in recent years to critical periods of development than to the phenomenon of poorer memory in infancy. Nonetheless, we believe that this is a phenomenon of major significance and one that deserves much more attention. As a first step, it is necessary to distinguish between memory of a past experience and other lasting effects of that experience.

Memory, in the view of most psychologists, is a loosely defined concept that describes the ability of animal and human subjects to recall or relearn more rapidly previously practiced behaviors. As concepts, learning and retention have defied precise definition, and it is well beyond the scope of the present review to attempt to improve on that state of affairs. Nonetheless, it is evident that much of what is influenced by early experience is mediated by changes in neural, endocrine, autonomic, and other structures; changes that can in no way be considered learned although they may drastically alter subsequent capacity to learn. Total light deprivation in infancy, for example, produces retinal degeneration (Riesen, 1961), and deprivation of patterned light alters the adult neurophysiological response to certain visual stimuli (Wiesel & Hubel, 1965). Similarly exposure to early stress permanently changes adrenal reactivity (Levine & Mullins, 1968), and rearing animals in enriched environments appears to alter brain size and neurochemical activity (Rosenzweig, Krech, Bennet, & Diamond, 1968).

In a like fashion, there are critical periods in development where exposure to members of the same species is necessary for the individual to acquire many species-specific behaviors. These critical interactions appear to be genetically preprogrammed so that normal development occurs only if the appropriate stimuli are present at predetermined times. While these experiences may permanently alter underlying neural or physiological structures, there is no evidence that the organism necessarily "remembers" the specific stimuli to which it was exposed. In man, a child may develop normally under the care of one parent for the first few years of life, but have no specific recollection of that parent if separated from it and subsequently reared in another family setting. Yet the parental care given during those first few years is critical for normal development. The ability to recall a specific event is often independent of the effects of that experience.

---

[1] This work was supported in part by National Institutes of Health Grants MH01562 and MH08501 to Byron A. Campbell and in part by National Institutes of Health Grants MH12064, MH18619, and MH47359 to Norman E. Spear.

[2] Requests for reprints should be sent to Byron A. Campbell, Department of Psychology, Princeton University, Princeton, New Jersey 08540.

It is the premise of this paper that learning and retention represent discrete psychological phenomena which can be investigated developmentally independent of other lasting effects of early experience, effects which we believe are mediated by processes other than simple retention of learned material. Moreover this paper will focus on phenomena which are customarily studied under the heading "long-term retention," with particular attention paid to changes in retention taking place over days, weeks, or even months. In addition, an effort will be made to distinguish between developmental differences in retention due to strength of original learning versus developmental differences in retention due to differences in retention processes per se; the present paper is addressed primarily to the latter.

I: DEVELOPMENT OF LONG-TERM MEMORY

Ontogenesis of Long-Term Memory in Man

In man, the rapid forgetting of information acquired during early childhood is often described as infantile amnesia (Schachtel, 1947). When adolescents, young adults, or the aged are asked to recall the events of their childhood it is universally evident that little of the first three to five years of life is subject to conscious retrieval (cf. Dudycha & Dudycha, 1941; Schachtel, 1947; Waldfogel, 1948, for comprehensive reviews). Most individuals are completely amnestic for that period being unable to recall even a single event occurring prior to three years of age and some have no childhood memories prior to age nine (Saul, Snyder, & Sheppard, 1956). The number of events than can be recalled for any one year also increases sharply with age (Waldfogel, 1948). Even then there is considerable question as to the nature of these early memories, since the most vivid memories of early childhood are those that are frequently recalled or rehearsed. For most young adults, recall of early childhood usually begins to achieve some clarity by the late preschool or early school years, with the conciseness of memory continuing to increase through the primary school years (Waldfogel, 1948). The majority of research on infantile amnesia took place prior to 1940 during which time numerous theories were put forth to explain the phenomenon. Allport (1937, p. 161) summarized the four major theories of that time as follows: (*a*) Infant experiences are not verbalized and, therefore, cannot be held as concepts in consciousness; (*b*) the areas of the cortex involved in conscious memory are not myelinized and, therefore, are unable to hold "traces" of experience; (*c*) the apperceptive context of early life is so different from that of adulthood that events occurring in it cannot be recalled; (*d*) when the infant encounters the "reality principle" he represses into the unconscious all memories of his self-indulgent life up to that date. It goes without saying that none of these explanations have been empirically established; instead the entire phenomenon of infantile amnesia appears to have been repressed by the scientific community rather than investigated systematically. One goal of the present paper is to reawaken interest in this important phenomenon.

Unfortunately infantile amnesia has not been the target of many well-controlled experimental studies. Levy (1960) in a study performed for other reasons, found that memory of innoculation increased from birth to 18 months of age. Before six months, infants showed no memory of previous injection experience; after that, retention increased gradually with age. These data offer some experimental support for the phenomenon of infantile amnesia, although they give no indication as the whether the rapid forgetting of infancy is due to poorer original learning or to a deficiency in the retention process per se. In general the phenomenon of infantile amnesia is in dramatic contrast to the extent and quality of learning that takes place during those years. Preschool children acquire a vast array of complex behaviors and show striking memory for preschool and home incidents over short intervals. Yet years later they are completely amnesic not only for specific incidents but for the entire environment as well.

The absence of careful experiments studying long-term forgetting by children compared to adults is no mystery. Practical barriers in such experiments are considerable, some of which have been noted in reviews by Keppel (1964) and Goulet (1968). The most immediate difficulties concern the task: Material to be learned must not be so simple as to be immediately overlearned by all adults but not so difficult as to be impossible or frustrating to learn by younger subjects. For example, words or numbers cannot be used with most preschoolers because they simply cannot read them.

Once a task is selected, the next step is to devise means for controlling the degree of original learning prior to initiating the retention interval. As Keppel (1964) and Goulet (1968) have noted, the failure to achieve this has made nearly all potentially relevant experiments conducted prior to 1960 of little value in analyzing changes in retention during development. Appropriate procedures have been devised for verbal learning (Underwood, 1964), but their application in experiments with children is time consuming and expensive in practice (e.g., Amster, Keppel, & Meyer, 1970).

In recent years an increasing number of well-controlled experiments on age and retention of verbal learning have appeared, but they have been directed more toward implications of interference theory than the development of memory itself. Thus, these studies have included either young subjects who are too old (e.g., Walen, 1970) or retention intervals that are too short (Bach & Underwood, 1970) to provide experimental evidence relevant to our concern with infantile amnesia.

Certain of the methodological problems encountered in studying the development of long-term memory with human subjects can be overcome when animals are used as subjects, a circumstance common to many areas of psychology and biology. Campbell (1967) has described some recent advances in methodology which permit clearer interpretation of the deficit in long-term retention found in young rats compared to adults. We turn now to a brief review of these experiments.

Ontogenesis of Long-Term Memory in Infraprimate Mammals

During the first half of this century there were numerous ontogenetic studies of learning as part of an effort to understand and define animal intelligence, but little attention was focused on long-term retention despite the then current interest in early childhood memory. The first major parametric ontogenetic investigation of long-term memory was reported by Campbell and Campbell (1962). In this study 18-, 23-, 38-, 54-, and 100-day-old rats were conditioned to fear a distinctive compartment of a double compartment shuttle box, and then tested for retention of fear 0, 7, 21, or 42 days after original training. Retention, as measured by avoidance of the fear side, increased dramatically with age. Eighteen-day-old rats showed nearly 100% avoidance of the fear side when tested immediately after training, but only chance avoidance at the 21- and 42-day retention intervals. In contrast, 54- and 100-day-old subjects showed little or no forgetting at any retention interval. In a second experiment, retention of conditioned punishment was investigated over 0-, 21-, and 42-day intervals in rats which were 23 and 100 days of age when conditioned. Again, juveniles were clearly inferior to adults.

Similar findings have been obtained with active avoidance (e.g., Kirby, 1963), escape (e.g., Smith, 1968), and appetitive discrimination (Campbell, Jaynes, & Misanin, 1968) tasks. Kirby trained 25-, 50-, and 100-day-old rats on a simple runway avoidance problem. The age groups did not differ during acquisition or extinction, but 25-day-old subjects showed inferior retention in relearning tests 25 and 50 days following original learning. Smith found no significant differences during acquisition of a T-maze escape task, yet performance of 25-day-old rats was substantially poorer following a 75-day retention interval than those of their adult counterparts who showed almost no forgetting. Campbell et al. trained young and adult rats to approximately equal performance levels on a discriminated operant task (lever press for food). Juveniles showed considerably poorer retention than adults of both the discrimination and the basic instrumental response in tests (extinction) 38, 75, and 150 days after original training.

Crockett and Nobel (1963) and Thompson, Koenigsberg, and Tennison (1965) have failed to find age differences in retention. In the Crockett and Noble experiment, relearning of a light-motivated escape response was comparable in rats 21 and 90 days of age after a 60-day retention interval. Unfortunately control groups without prior acquisition experience were not included, making it difficult to interpret the results. Thompson et al. trained 25- and 100-day-old rats in avoidance shuttle boxes scaled to the physical size of the animals. No differences in retention were obtained when the physical size of the apparatus was increased proportionately at the time of relearning, but no young subjects were trained and tested in the same size apparatus. According to the generalization decrement hypothesis (cf. pp. 222–224), greater forgetting would have occurred under this condition.

More recently, Schulenburg, Riccio, and Stikes (1971) and Feigley and Spear (1970) have shown in several experiments that retention of passive-avoidance responding also increases with age. Similarly, Campbell, Lytle, and White[3] have systematically compared acquisition and retention of escape and passive avoidance responding in preweanling and weanling (15–35-day-old) rats. In these extensive experiments, rate of forgetting was inversely proportional to age, with the 15-day-old subjects showing extremely rapid forgetting.

Strength of Original Learning versus Differences in Retention and Retrieval Mechanisms as Determinants of the Rapid Forgetting in Infancy

Taken together, the above studies demonstrate unequivocally that rate of forgetting is more rapid in infancy than adulthood. But what is not clear is whether the rapid forgetting of infancy is due to differences in strength of original learning or to ontogenetic differences in the storage and/or retrieval mechanisms of retention. Any of these could underly the observed differences in long-term retention.

There are, however, a number of features of the preceding studies which suggest that the rapid forgetting of infancy represents a deficit in retention and is not simply the result of poorer original learning. One critical question is whether young rats forget more rapidly than adults when degree of original learning is equal to or greater than that of the adult. Several different techniques have been used to assess and equate degree of original learning in ontogenetic studies of memory in animals. They are as follows:

Rate of acquisition. Concerning the problem of assessing retention effects unconfounded by learning differences, Underwood (e.g., 1964) has noted that final degree of learning will usually be lower following slower rates of attaining a common criterion of performance. However, it is usually assumed (although there are exceptions) that strength of original learning is the same for all groups attaining a common criterion of performance prior to a retention test if rates of attaining that criterion are the same. Several developmental studies appear to meet these conditions. As noted earlier, the studies of Kirby (1963), of Feigley and Spear (1970; active avoidance), and of Smith (1968; T-maze escape) found no systematic differences in either the rate or asymptotic level of learning, yet forgetting was significantly greater in the younger subjects. Similarly in the Campbell et al. (1972) study, rate of forgetting in the T-maze escape experiment was more rapid in 20-day-old rats than in the 25- and 35-day-olds even though rate of acquisition was at least as rapid in the younger age group. In the passive-avoidance experiment by Campbell et al., there were no apparent differences in the number of trials to criterion between the 16- and 25-day-old animals, yet the 16-day-old subjects forgot significantly faster. In addition there are a number of instances where the retention of older subjects is superior even though rate of learning was significantly slower or poorer in those groups (e.g., the 100-day-old subjects in the Kirby experiment and the 35-day-old subjects in the Campbell et al. study [see Footnote 3]). Thus when degree of original learning is assessed by rate and asymptotic level of learning, retention of equally learned responses remains typically poorer in young rats.

---

[3] B. A. Campbell, L. D. Lytle, and B. C. White. Species differences in ontogeny of memory: Support for neural maturation as a determinant of forgetting. Submitted to the *Journal of Comparative and Physiological Psychology*, 1972.

Resistance to extinction. Another widely used measure of response strength is resistance to extinction. By this measure, strength of original learning was the same across ages in the Campbell and Campbell (1962) and Kirby (1963) experiments. As noted before, forgetting was more rapid in the younger subjects.

Effect of overtraining. Another approach is to deliberately overtrain young subjects, by administering either more intense shocks or longer durations of shock. Using this procedure, Schulenburg et al. (1971) found that lengthening shock duration from one to five seconds increased long-term memory in 15- and 21-day-old rats on a passive-avoidance task, but that these young subjects were still markedly inferior to adult subjects receiving only one second of shock. Using a similar approach, Feigley and Spear (1970) varied intensity of shock during acquisition of a passive-avoidance response. Weanling subjects showed inferior retention at all shock levels when compared to adults.

Relearning. Another indication that strength of original learning is comparable across ages in the preceding experiments stems from the finding that the number of trials required to relearn the criterion response does not differ at short retention intervals. Klein and Spear (1969), for example, compared weanling and adult rats on the relearning of an avoidance response .08, 1.0, 4.0, or 24 hours after original learning and found identical retention functions ("Kamin Effects") for both groups. Long-term retention of avoidance responding is, as noted earlier, much poorer in weanlings than adults. This and the above findings demonstrate that at least some responses can be learned and stored for short intervals with the same fidelity in young and adult animals and suggest that other processes specific to the early stages of development act to produce the poor long-term memory characteristic of infancy. This is not to imply that all responses are acquired equally at all ages nor that much of what is subsumed under the phenomena of infantile amnesia in man is not the result of weaker original learning. We do, however, wish to emphasize that forgetting of responses acquired in infancy is more rapid than those acquired in adulthood, and that this rapid forgetting is not necessarily due to poorer original learning.

In this context, the aim of the present paper is to examine some of the processes most likely to underly the ontogenesis of long-term memory. In our view, two broad classes of variables are potential determinants of the poorer memory found in infancy. First there are a number of interference and information-processing mechanisms which may change with age. These include retroactive and proactive interference, stimulus generalization decrement, stimulus differentiation, and related processes, all of which might act to impair long-term memory more in the early stages of development. Second, there are a variety of structural and chemical changes in the central nervous system occurring during this period which may determine the changing memory capacity of the developing animal. These variables are not necessarily independent, since maturation of the central nervous system may alter the shape of the primary stimulus generalization gradient, susceptibility to interference, etc. Nevertheless, our goal is to examine the relative importance of these two classes of variables as determinants of poorer memory in infancy.

As will soon become apparent, there is little evidence to support either view. Hence the major purpose of the present paper is not to deduce which of the many possible mechanisms underlies the rapid forgetting of infancy. Instead the purpose is to describe in some detail a variety of processes that may underly this change. These possibilities stem directly from present knowledge of behavioral processes and neurological development. Our prime hope in writing this paper is to stimulate research on what we consider a major problem of unusual current importance.

II: DEVELOPMENTAL CHANGES IN INTERFERENCE AND INFORMATION-PROCESSING MECHANISMS

Retroactive and Proactive Interference as Determinants of Poorer Memory in Infancy

Until the late 1950s, it was generally agreed that retroactive interference was the major source of forgetting (e.g., McGeoch & Irion, 1952). But in 1957, Underwood demonstrated that forgetting in the laboratory was directly proportional to the number of previous lists learned, a finding which ultimately led many investigators to conclude that proactive interference was the major source of forgetting, both in the laboratory and in normal life (e.g., Underwood & Postman, 1960). This emphasis also stimulated a search for extraexperimental sources of interference, particularly within the context of language. While the results of this research are not unambiguous, it is generally accepted that extraexperimental experiences can contribute substantially to the proactive interference pool (e.g., Postman, 1971).

In the context of developmental psychology, the question is a simple but unanswered one: What are the relative contributions of retroactive and proactive interference to the basic phenomenon of poorer memory in infancy? According to the simplest model of proactive interference, older animals should forget more rapidly because they have lived longer with a correspondingly higher probability of having acquired previously learned habits which might interfere with retention of more recently learned material. But as the data cited above indicate, this prediction is contrary to fact.

Retroactive interference, in contrast, remains a somewhat more viable candidate as the basis for the rapid forgetting of infancy. This would be particularly true if the total amount learned at any developmental stage were proportional to the animal's overall level of behavioral activity. Figure 1 illustrates how exploratory behavior, as measured by a four-hour exposure to a simple stabilimeter cage scaled to the size of the rat, changes during development (from Moorcroft, Lytle, & Campbell, 1971). As can be seen, little activity occurs prior to 10 days of age, but between 10 and 15 days of age activity leaps to a level 3–4 times that of the 10-day-old and nearly 10 times that of the adult rat. Hyperreactivity to novel stimuli and generally high levels of exploratory and play activity also rise and then decline in the normal home cage environment during these same ages (Bolles & Woods, 1964). The decrease in overall activity levels parallels, from 15 days of age to adulthood, the development of long-term memory. While it is not typical to assume a correlation between activity level and learning, it is not an unreasonable possibility. One could easily argue that the young animal is learning a great deal about the spatial-perceptual-social characteristics of the environment as it moves about; the more it approaches and withdraws from stimuli in different directions and under varying conditions, the more it learns. In contrast, the older animal, inactive and maintained in a familiar environment, learns less than its weanling counterpart. Given this analysis, which is obviously not dependent on gross levels of motor activity although supported by those findings, it follows that retroactive interference from learning occurring during the retention interval is likely to be higher in young than adult animals.

On the other hand, the implicit assumption stemming from the data presented in the previous section is that long-term memory increases monotonically between birth and maturity. Thus the low level of activity prior to 10 days of age makes retroactive extraexperimental interference an unlikely basis for the presumably rapid rate of forgetting during this period. However, no developmental retention studies have been reported showing age-related differences in retention unconfounded by learning differences among rats younger than 15 days. Separation of learning and retention factors among such neonates probably will prove most difficult, if possible at all. What might stem from this pattern of behavioral development is a two-factor theory of forgetting during ontogenesis. During the early stages of development, long-term forgetting may be due primarily to immaturity of the central nervous system, accompanied by closely linked deficits in short-term memory and learning rate; later in development interference-type processes together with the consequences of central nervous system growth (see Section III) may play a greater role in determining rate of long-term forgetting.

It is also impossible to distinguish between retroactive and proactive interference on the basis of these data. First, proactive interference is known to be more effective when acquired immediately prior to original learning (Spear, 1971; Underwood & Freund, 1968). In the present instance, rate of forgetting appears to be correlated with the amount of exploratory activity elicited just before original learning as well as the amount of interpolated activity. Second, proactive interference produces less interference when prior learning is acquired through spaced learning over long periods of time or, more generally, when the interfering and critical materials are easily differentiated (Underwood & Ekstrand, 1966). In the present instance it seems not implausible to assume that extraexperimental proactive interference is minimized for the adult by the extensive familiarity subjects acquire with proximal environmental stimuli during the course of development. To the extent that the rapid forgetting typical of early developmental stages is due to interference processes, it seems difficult if not impossible to assess the relative contributions of proactive and retroactive interference to this process. In all likelihood, both contribute substantially to memory decrements over time, with their relative weights being determined by the specific experimental or situational circumstances. Moreover, as Deese and Hulse (1967) point out, the theoretical distinction between retroactive and proactive interference is greatly diminished when long retention intervals are used.

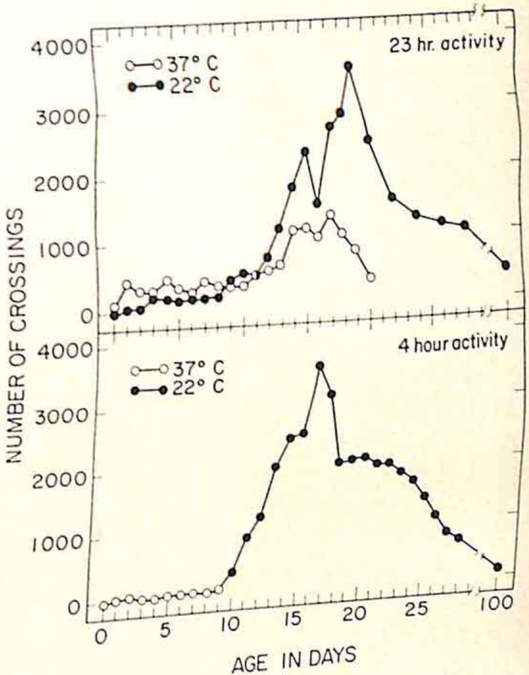


FIG. 1. Spontaneous activity of rats in stabilimeter cages during postnatal development. (Each point in the upper panel represents a different group of six to eight animals tested for 23 hours, while the data in the lower panel represent a single group of animals tested for four hours each day.) (Reprinted from an article by W. H. Moorcroft, L. D. Lytle, and B. A. Campbell published in the *Journal of Comparative and Physiological Psychology*, 1971, Vol. 75. Copyrighted by the American Psychological Association, 1971.)

Maturation of Species–Specific Behavior

During the normal course of development, many drastic changes take place in the animals' behavioral repertoire. Some are learned, some are contingent upon changes in physical capability, some are preprogrammed in unknown ways such that they emerge at a particular stage when the proper stimulus appears, and still others are hormonally dependent. Male preweanling rats respond to the dam by rooting and sucking; at sexual maturity the same rat responds to the same female by copulating. Similarly, pain produces withdrawal and flight in the young animal and attack in the adult (Hutchinson, Ulrich, & Arzin, 1965), and unfamiliar environments may elicit more fear or general arousal in older animals (Candland & Campbell, 1962; Sluckin, 1965). These age-related modifications of behavior could act to either permanently or temporarily interfere with the performance of a previously learned response as an animal develops. As a hypothetical example, we suspect it would be possible to train an immature female ring dove to peck at a ring dove egg in a nest for food reinforcement, and that the dove would show good retention of the response over long time periods. But if the bird was brought into broodiness by parturition or other means, it seems quite likely that the bird would now respond to the eggs by brooding rather than pecking, resulting in a measured loss in retention of the pecking response. In this dominance of unlearned behavior over learned behavior did develop, it would then be of further interest to see if the learned behavior reappeared when the hormonal or other conditions controlling the unlearned behavior were removed.

In short, there must be many situations in which an emerging unlearned response takes precedence over a previously learned response to that stimulus. The extent of this interference is undoubtedly proportional to the animal's repertoire of unlearned behaviors and is thus unlikely to be a major source of forgetting in man.

Maturational Changes in Information-Processing Mechanisms

As a corollary to the above observation that emerging of species-specific behaviors may interfere with previously learned responses, it seems reasonable to consider the possibility that changes in mode or style of information processing may also interfere with the retrieval of previously learned material. In higher mammals such as man, enormous changes take place in the way in which sensory information is encoded and responded to. The most obvious changes occur in the use of language. To a preverbal child, a book is simply another object in its environment; later the object is verbally labeled as "book"; still later it is responded to by reading either overtly or covertly the title; and finally the title evokes a variety of emotional or evaluative responses ranging from interest to aversion. Underwood (1969, p. 571) has suggested other instances in which attributes forming critical components of childhood memories may not correspond to those forming comparable memories in the adult, and at least one subsequent experiment has yielded confirming evidence (Bach & Underwood, 1970). Failure to retrieve memories acquired prior to maturity might be one consequence of this shift in mode of information processing. Similarly modes of problem solving appear to change developmentally in man from overt stimulus–response solutions to covert mediated solutions utilizing a complex of previously acquired behaviors (Kendler & Kendler, 1962).

Findings and observations such as these again raise the question as to the effect of such changes on memory. Is memory for a task learned at an early stage of development less accessible because stimulus information of that type is now processed in a different fashion? This is an intriguing possibility because the greatest changes in information processing appear to take place in preschool and early school years, exactly those periods for which typical adults have the poorest recall (cf. Inhelder, 1969). To date there have been few empirical investigations on the effects of changes in mode of information processing, but we believe it to be a promising area for research even though the methodological difficulties are formidable.

Differentiation

During the course of development, all mammals must pass through a stage in which they actively learn the stimulus characteristics of their environment. The motivation and reinforcement for this learning are obscure and are the source of continuing discussion, but whether motivated by curiosity, fear, or primary or derived motives and incentives, it is irrelevant to the observation that such learning does take place. Even the lowly laboratory rat, housed in stimulus-impoverished environment, learns the characteristics of his home cage environments (Candland & Campbell, 1962). The implications of this observation for our analysis of developmental changes in memory are substantial; if an animal does not differentiate between the home cage environment and the test environment, criterion responses acquired in the latter will either be enhanced or extinguished in the home environment depending on the similarity of stimulus–response contingencies in the two environments.

Stimulus Generalization

By definition, stimulus events are never reproduced completely; consequently, performance on each successive trial in a learning task, regardless of intertrial interval, is determined by the perceived similarity of the ongoing trial to previous trials and the strength of the acquired response. In well-trained adult animals, presumably little stimulus generalization decrement occurs from trial to trial even over long retention intervals, but in young animals, the possibility exists that growth-induced changes in body and physiological makeup substantially alter the perception of the test environment resulting in a performance decrement. A description of some of the growth-induced changes which might act to impair performance of habits acquired early in development follows.

Changes in perception of the conditioned stimulus. By far the most commonly considered source of age-induced stimulus generalization decrement is the change in perceived size of the environment that is thought to occur as the animal matures. From this it follows that to a young rat, a maze may look like a great cavern from which escape requires a massive effort, but on retest after a retention interval, the same apparatus may appear much different because the visual angles from which the apparatus is viewed have changed or because the size of the apparatus, relative to the animal's body size, is now much smaller. Similarly the child's view of the world changes from one which is populated by giants and distinguished by spots of gum underneath tabletops, to the adult view of the world as we know it. This hypothesis is so intuitively appealing that some psychologists have considered this the major mechanism underlying forgetting in infancy. Perkins (1965), for example, states in reference to Campbell and Campbell's (1962) report showing poorer retention of fear in infancy that "the stimulus situation to which fear was learned may be so altered by stimulus changes resulting from growth that there is nearly complete generalization decrement resulting from maturation induced changes [pp. 46–47]." Additional evidence for this view stems from the observation that handwriting characteristics of the early school years reappear when adult humans are asked to write their names slowly, and that long-forgotten memories are sometimes recalled when phrases typical of childhood such as "No, I won't" are shouted in the uncontrolled fashion of childhood (Tomkins, 1963, p. 334).

Maturational changes in perception of unconditioned stimuli. When relearning is used as a measure of retention, the possibility exists that the unconditioned stimulus will be perceived as either qualitatively or quantitatively different. For example, a specified intensity of shock may produce intense pain in young rats and only mild stimulation in adult rats resulting in a performance decrement. Research on this particular question is ambiguous. Campbell (1967), using a matched-impedance shock source, found no differences in aversion thresholds across a wide range of ages in either rats or guinea pigs, but Paré (1969) has recently reported that young rats are much more sensitive to constant current dc shock than adults. Similarly Klein and Spear (1969) have obtained data suggesting that sensitivity to electric shock changes between 15 and 25 days of age, and Campbell (1967) observed that the tetanization thresholds were lower for young animals. This factor probably does not seriously affect the data showing poorer retention by infants than adults, however. Feigley and Spear (1970) have shown that the extent to which weanlings (21–25 days old) show greater retention deficit than adults does not differ over wide variations in shock intensity for both active- and passive-avoidance tasks. Furthermore, the magnitude of the retention deficit was not influenced by shock intensity for either weanlings or adults.

It is also difficult to equate amount of reinforcement in appetitive studies across ages. Obviously the typical 45-milligram Noyes pellet constitutes a greater proportion of the young animal's nutritive requirements than the adult animal's, and the difficulties in scaling the size of the pellet during retesting constitute major methodological questions (cf. Campbell, 1967; Campbell & Jaynes, 1968). Whether changes in absolute pellet size or changes in pellet size relative to body size produce generalization decrement has not been investigated.

Experimental tests of the generalization decrement hypothesis. To date there have been few experimental tests of the generalization decrement hypothesis, and the results are thus far ambiguous. In the first direct test, Thompson et al. (1965) trained 25-, 50-, and 100-day-old rats in shuttle boxes scaled to the size of the rat, and then retested them in the largest of the three boxes after retention intervals of 0, 50, or 100 days. Using this procedure, no differences in long-term retention were found as a function of the age at time of training, a result which they interpreted as showing that the rapid rate of forgetting typically found in young rats is the result of generalization decrement. Unfortunately the design of the experiment is incomplete in first failing to demonstrate rapid forgetting in young rats with test apparatuses of the same size, and then in not demonstrating that adult animals show a generalization decrement when apparatus size is changed in the retention test. In addition, rate of acquisition varied significantly across age groups, further complicating interpretation of the results.

In a more recent test of the generalization decrement hypothesis, Feigley and Spear (1970) trained young rats in an avoidance box appropriate for adults or in a smaller apparatus roughly scaled to the weanling body size. After a 28-day retention interval they were tested in the larger box; thus, rats trained in the smaller box were tested in an apparatus which increased in volume by a magnitude roughly proportional to the rats' own increase in body size. No differences in this long-term retention were found, thus contradicting the conclusions of Thompson et al. With both boxes, Feigley and Spear found retention of the young rats to be significantly poorer after 28 days than adult rats trained and retested in the same-size apparatus.

Stimulus generalization gradients as a function of age. A question closely related to the above issue of whether or not young animals show growth-induced generalization decrements during the retention interval concerns the nature of the primary stimulus generalization gradient itself. Many theorists (e.g., Prokasy & Hall, 1963) have contended that the primary stimulus generalization gradient reflects the animals' failure to discriminate the original training stimulus from the test stimulus. From this reasoning it follows that young animals with little prior stimulus experience should show flatter generalization gradients than adults. Unfortunately, it is not clear at this time whether generalization gradients in infancy are typically flatter, steeper, or the same as in adulthood, since there is conflicting evidence to support each view (Frieman, Rohrbaugh, & Riccio, 1969; Peterson, 1962; Rudolf, Honig, & Gerry, 1969). In the present context, flatter generalization gradients in infancy may leave the younger animal more susceptible to extraexperimental interference or facilitation, while steeper generalization gradients in infancy may maximize decrements in response strength produced by maturation-induced changes in perception.

III. NEUROLOGICAL CORRELATES OF MEMORY DEVELOPMENT

In addition to above described behavioral changes taking place during development that might be related to the poorer memory of infancy, there are an equal or greater number of changes in the central nervous system which might also influence memory. The purpose of this section is to review some of those changes and to speculate on the manner in which they might produce the observed ontogenetic differences in memory. Before proceeding to this discussion, it is necessary to reiterate the distinction between the processes of learning and memory, since learning capacity may mature at one rate and retention capability at another. In the data reviewed earlier, there were numerous instances in which young animals acquired new material at the same rate as adult animals but then showed much more rapid loss of those responses over long time periods. Within the field of human memory, for example, it is well established that certain variables (meaningfulness, intralist similarity, intertrial interval) may have quite different effects on learning rate and retention (cf. Keppel, 1964; Spear, 1970; Underwood, 1966), thus providing further precedent for instances of functional independence between learning and retention.

The ontogenetic distinction between learning and retention recalls the concept of encephalization as Yerkes (1912) applied it to a study of retention in the earthworm. In a classic experiment, Yerkes trained earthworms to avoid shock and approach moist earth in a T-maze. At criterion, the caudal section was severed from the head ganglion and then retested periodically in the maze. Surprisingly, the caudal section appeared to retain the response until a new head ganglion regenerated. Yerkes interpreted this as showing that memory could be stored in the worms' primitive ladder nervous network, but that regeneration of the head ganglion (encephalization) somehow dominated or disrupted the memory trace resulting in loss of the habit. Unfortunately this experiment has never been successfully replicated, and it now seems likely that Yerkes' original finding was an artifact (Ratner, 1967).

Nonetheless, the concept of encephalization-induced memory loss is one which may have considerable relevance to the high rates of forgetting observed in young mammals born with immature central nervous systems. In the simplest sense, it seems quite possible that continued neurological development following the acquisition of a learned response could act to either disrupt the memory-storage mechanism(s) or to block retrieval of previously learned materials. Surprisingly, there has been little or no speculation about either of these possibilities during the past several decades or more, a period during which there has been an astronomical increase in the amount known about the central nervous system and its development.

Myelinization

Prior to the 1950s, functional development of the central nervous system was most commonly linked to myelinization. Flechsig (1896) was the first to observe that myelin was either absent or grossly deficient in neonatal mammals, a finding which led him to postulate that myelinization preceded function. In man, for example, he believed that the ability to form associations did not appear until two to three months postnatally, the time at which myelin is first observed in the cortical association areas. Elaborating on this reasoning, many early psychologists attributed the phenomenon of infantile amnesia to incomplete myelinization (Allport, 1937, p. 161). In fact, there is a striking correlation between the development of long-term memory in both rat and man and myelinization.

In mammals, myelinization proceeds in a roughly caudocranial order. Myelin first appears in the spinal cord, then in the medulla, pons, and midbrain, and finally in the telencephalon. The cerebral cortex also myelinates sequentially starting at the projection centers and ending at the association centers (Jacobsen, 1963).

In the rat, the first evidence of myelin appears in the ventral spinal roots at about 2 days postnatally, and it is not until the rat is about 10 days old that myelin begins to appear in the cerebrum. Within the telencephalon, thalamic fibers begin to myelinate between 10 and 12 days and are not fully myelinated until 40–50 days postnatally. In the cortex, myelinization does not begin in many regions until 21 days and is not complete until the rat is approximately 60 days of age (Jacobsen, 1963) with the period of most rapid deposition occurring between 30 and 40 days (Bass, Netsky, & Young, 1969).

The major effect of myelin in nervous tissue is to increase speed of nerve conduction. The thicker the myelin sheath, the more rapid the speed of transmission. Consequently the changes in speed of neural conduction that occur developmentally are nothing short of remarkable. Figure 2 illustrates the increase in conduction velocity that occurs in the corticospinal tract of the kitten. Comparable changes occur in other brain regions and in other mammals, such as the rat and man, which show comparable changes in myelinization (Grafstein, 1964; O'Brien, 1970).

A related characteristic of the young organism is the relative fatigability of its central nervous system. With repetitive stimulation, the refractory period of cortical responses increases and the amplitude decreases relative to the adult pattern of responding. As a result of these factors, the central nervous system of the young animal is only able to receive and treat a relatively small amount of information per unit of time (Scherrer, 1968, p. 488).

How might these changes affect memory? Clearly deposition of myelin in axon sheaths is not a primary memory storage mechanism, nor is it likely that changes in the refractory period directly influence memory, yet these changes may interfere with retrieval of learned material by a simple principle analogous to primary stimulus generalization.

If the central nervous system can process only limited amounts of information early in development, then the amount of information encoded in a stimulus sequence preceding the elicitation of a criterion response is necessarily less in infancy than adulthood. Consequently, when an animal is tested for retention after a period during which neurological growth has occurred, the amount of information processed per unit time in the test environment would be greater than it was during original learning. This change could conceivably result in a generalization decrement leading to poorer performance. Similarly changes in speed of transmission per se could produce a generalization decrement. Proprioceptive feedback, for example, is more rapid in the adult than young rat.

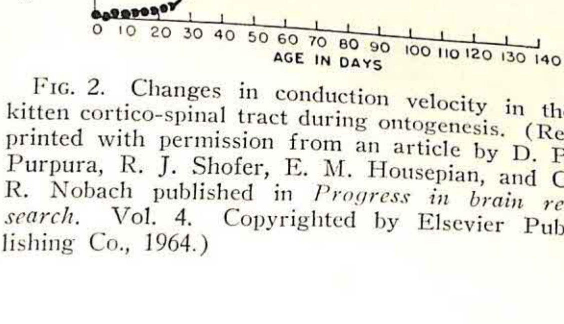


FIG. 2. Changes in conduction velocity in the kitten cortico-spinal tract during ontogenesis. (Reprinted with permission from an article by D. P. Purpura, R. J. Shofer, E. M. Housepian, and C. R. Nobach published in *Progress in brain research*. Vol. 4. Copyrighted by Elsevier Publishing Co., 1964.)

Cell Differentiation

Another dimension of central nervous system growth that may influence learning and memory is the maturation of cell structure. In immature mammals, the pyramidal cells of the cortex, for example, are small, densely packed, and the dendritic structure incompletely elaborated (Eayrs & Goodhead, 1959; Purpura et al., 1964). As maturation proceeds, the dendritic structure becomes more and more complex until the cell acquires its adult characteristics. The time course of this growth sequence in the rat has been illustrated by Eayrs and Goodhead (1959) and is reproduced in Figure 3.

Using a Golgi stain and a light microscope technique, Eayrs and Goodhead also estimated the growth of axodendritic junctions by counting the number of times dendrites of a given cell appeared to make contact with axons. Their findings are summarized in the left panel of Figure 4. More recently, Aghajanian and Bloom (1967) counted the number of synaptic junctions in the molecular layer of the rat's parietal cortex at different stages of development using electron microscopy. Their results are reproduced in the right panel of Figure 4. In both cases it is evident that the majority of synaptic junctions are formed between 10 and 30 days of age, and it is just in this age range where the rat shows the greatest increase in long-term memory.

There are at least two ways in which number of synaptic connections may determine the poorer memory of the rat early in development. First it may be that memory is simply proportional to the number of

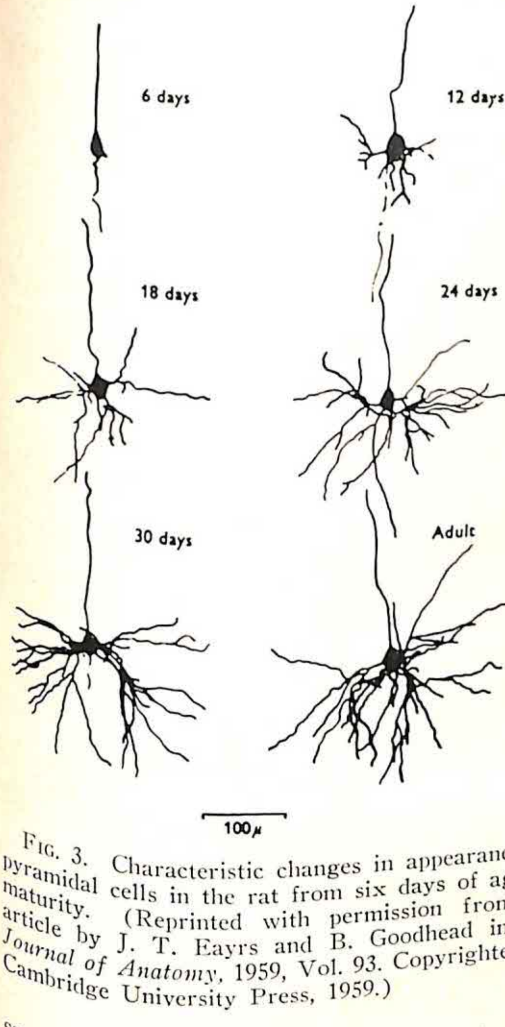


FIG. 3. Characteristic changes in appearance of pyramidal cells in the rat from six days of age to maturity. (Reprinted with permission from an article by J. T. Eayrs and B. Goodhead in the *Journal of Anatomy*, 1959, Vol. 93. Copyrighted by Cambridge University Press, 1959.)

synaptic junctions functioning at the time of original learning. Long-term memory may depend on redundancy. Second is the possibility that the increase in number of synaptic junctions following learning somehow impedes access to the previously learned response. Because of the vast increase in synaptic connections, sensory information necessary to elicit the learned response would almost surely be processed in different ways at different stages of development, thereby altering retrieval pathways. In short, it is quite possible that memories acquired early in development are stored with the same fidelity as adult memories, but that access to them is impeded by further central nervous system development. It is tantalizing to speculate that the reason early childhood memories return during senescence is because the most recently formed parts of the central nervous system are the first to atrophy (Paulson & Gottlieb, 1968).

In a more general sense, forgetting in infancy may parallel the disappearance of reflexive behaviors during ontogenesis. It is well established that the Babinski, rooting, and grasping reflexes, as well as many others, are inhibited by the functional maturation of higher brain centers, and that many of these reflexes reappear following cortical insult or arteriosclerosis. Thus the reflexes of infancy are always present in the central nervous system but inaccessible (not elicitable) for most of the typical mammals' life span. Because of the many parallels between reflexive and learned behaviors, it would not be at all surprising if some portion of what is learned in infancy is permanently stored but rendered inaccessible by further central nervous system development.

Development of Electrical Activity

Ever since Berger (cf. Brazier, 1961) succeeded in recording the electrical (electroencephalogram [EEG]) activity of the cortex, scientists have been attempting to identify components of this activity that might be associated with learning. To date this effort, while successful in many ways, has not identified the specific aspects of EEG

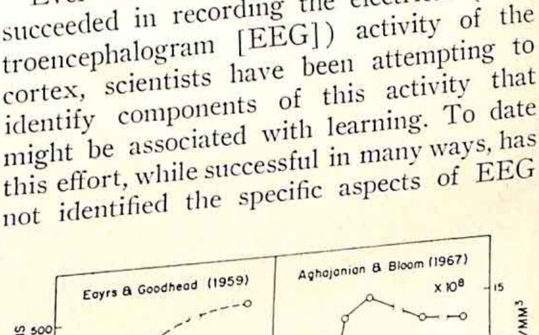


FIG. 4. Estimated number of synaptic junctions in the rat cerebral cortex during postnatal maturation. (Reprinted and adapted with permission from articles by the following authors: J. T. Eayrs and B. Goodhead in the *Journal of Anatomy*, 1959, Vol. 93. Copyrighted by Cambridge University Press, 1959; and G. K. Aghajanian and F. E. Bloom in *Brain Research*, 1967, Vol. 6. Copyrighted by Elsevier Publishing Co., 1967.)

activity that are critically associated with learning. Moreover, the extensive findings showing that memory persists following temporary disruption or cessation of EEG activity have led most investigators to abandon the notion that memory was stored in ongoing central nervous system electrical activity. Instead, most investigators now appear to believe that EEG activity is the manifestation of neurochemical activity at the synapse and that memory is encoded by morphological changes (e.g., chemical changes, neural growth, or changes in nucleic acid structure) either in the cell body or in the synaptic structure.

As a result, EEG activity can probably be best considered as a composite index of all brain components which influence nervous transmission. As such, the EEG becomes a useful index of maturational development of brain function. For the rat, there are a few signs of electrical activity in the cortex at birth. Action potentials are absent, although they can be elicited in subcortical structures, and there are few signs of spontaneous electrical activity. On Day 6, a primitive EEG pattern appears, but the major changes in the spectrum of EEG activity occur between 15 and 28 days of age. This progression of development is beautifully summarized in Figure 5 taken from Deza and Eidelberg (1967). Thus, as with many of the previous indexes of neurological maturation, there is a striking correlation between maturation of cortical electrical activity and the development of memory.

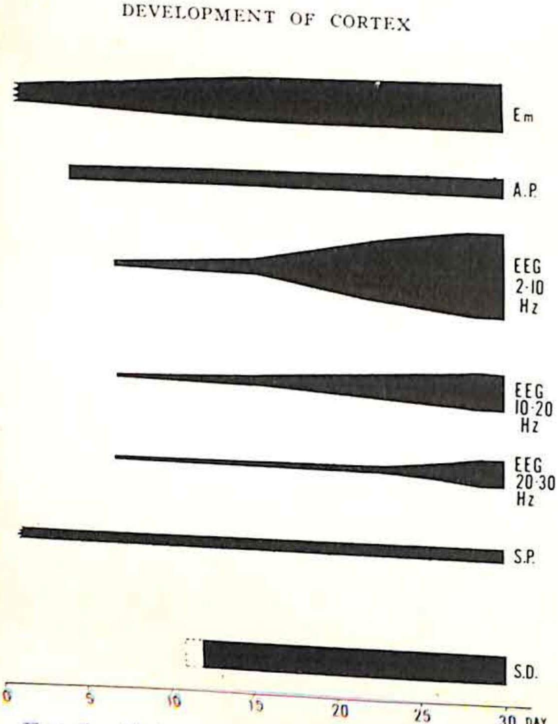


FIG. 5. Ontogenesis of electrical activity in the rat brain. (The width of the horizontal bars indicates progressive changes in either amplitude or frequency of brain electrical activity. The various measures are as follows: membrane potential [Em], action potential discharge [AP], cortical EEG activity [EEG], cortical steady potential [SP], and spreading depression [SD]. (Reprinted with permission from an article by L. Deza and E. Eidelberg in *Experimental Neurology*, 1967, Vol. 17. Copyrighted by Academic Press, 1967.)

Neural Transmitters

Of the known neural transmitters, acetylcholine is the most likely to be directly involved in the memory storage and retrieval process. Rosenzweig, Bennett, Diamond, Wu, Slagle, and Saffram (1969) have shown that enriched early experience, with the increased opportunities for learning that accompanies it, leads to an increase in both brain size and concentration of acetylcholinesterase, the enzyme which degrades acetylcholine. More directly relevant to memory and retrieval are the numerous studies showing that anticholinergic drugs such as scopolamine produce amnesia and/or dissociation (Berger & Stein, 1969a, 1969b). Conversely, anticholinesterases have been reported to enhance memory in some circumstances and impair it under others (Deutsch & Deutsch, 1966). The status and interpretation of this research remains very much in doubt, however, and there is as yet no consensus on the role that acetylcholine plays in the memory process. Among the other known central nervous system transmitters—norepinephrine, dopamine, and serotonin—none have been found to affect the memory process directly, although they may act to influence memory storage and retrieval by altering levels of arousal or emotional state (Seiden & Peterson, 1968; Stein & Berger, 1969).

Developmentally there is a remarkable similarity in the ontogeny of the major neural transmitters. All develop in a caudal-rostral direction (Friede, 1966; Robinson, 1968), and all show the greatest increase in concentration between 10 and 30 days of age, with more gradual increases continuing until the rat is 50–100 days old (for review see Friede, 1966; Himwich, 1970). Within the monoamine group, serotonin is the first to reach adult levels of concentration, norepinephrine second, and dopamine the last. Acetycholinesterase, which is probably correlated with concentration of acetycholine, also continues to increase until the rat is approximately 100 days of age (Friede, 1966).

This pattern of neural transmitter development parallels the development of memory, but again it is difficult to assign any special significance to the correlation between the two processes.

Deoxyribonucleic Acid, Ribonucleic Acid, Protein Synthesis, and Metabolic Rate

In recent years, numerous attempts have been made to implicate deoxyribonucleic acid (DNA), ribonucleic acid (RNA), protein synthesis, and metabolic rate in the brain to the memory storage and retrieval processes. While the status of these theories is unclear, a brief review of how these characteristics change during development in relation to the ontogenesis of learning and memory is in order.

Brain DNA and RNA both continue to increase postnatally in most mammals. Figure 6 shows the time course of these changes in several representative species. Increases occur in the dog, cat, rabbit, and rat; but no increase is seen in the guinea pig which is born with a mature central nervous system after a long gestation period. In the rat, both RNA and DNA attain their maxima about 14 days postpartum, well before long-term memory reaches adult levels, indicating that there is no simple relationship between total quantities of these brain constituents and memory. With respect to total DNA, an index of the number of cells in the brain, this is not surprising, since the mere presence of a cell body does not indicate that it has reached functional maturity. As noted earlier, dendritic arborization and the formation of synaptic clefts continues to develop long after the cell body is formed. In addition, the brain cells that are formed postnatally consist almost entirely of primitive microneurons, cells that are unlikely to be critical to learning and memory (Altman, 1970).

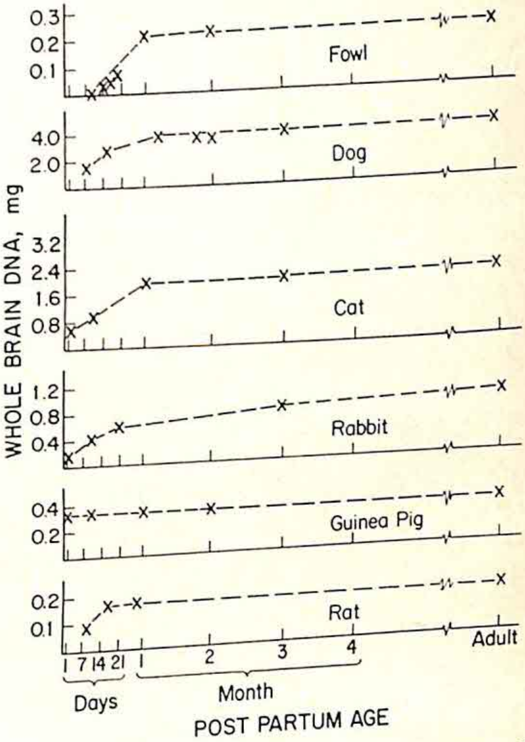


Fig. 6. Changes in total brain DNA content of several representative species during postnatal growth. (Reprinted with permission from an article by P. Mandell, H. Rein, Harth-Edel, and R. Mardell in *Comparative Neurochemistry*. Copyrighted by Pergamon Press, 1964.)

From these data, it also appears that total brain RNA is unrelated to the increase in long-term memory that occurs in the rat after 14 days postpartum. This is not to say that DNA or RNA are not essential substrates of the memory process, as they undoubtedly are (Booth, 1967), but to indicate that total quantities of these substances in the brain are simply not predictive of long-term memory.

In recent years, more attention has been paid to the overall role of protein synthesis in memory storage than to RNA and DNA. Numerous studies have shown that inhibition of protein synthesis by antibiotics such as puromycin and acetoxycycloheximide act to impair memory consolidation (Agranoff, 1970; Barondes, 1970; Flexner, Flexner, & Roberts, 1967). While there are many difficulties in interpreting these findings, there is a strong possibility that some aspect of protein synthesis is critical to memory. Unfortunately, the present synthesis inhibitors are ubiquitous in that they inhibit the formation of all proteins. This makes it virtually impossible to examine the relation between rate of protein synthesis and the development of long-term memory, since growth of the brain is predominantly dependent on synthesis of new proteins. When and if the proteins critical to memory storage are identified, then it may be possible to relate concentrations of those proteins to the development of long-term memory. At present, what is known is that the total amount of protein in the brain increases extremely rapidly beginning at age 10 in the rat and reaches an asymptote about 30 days of age (Agrawal & Himwich, 1970).

Another dimension of brain function that has been occasionally implicated in long-term retention is that of metabolic rate. Here the assumption is that the memory trace (or access to that trace) is gradually eroded over time by the turnover of cell body constituents, and that alteration in metabolic rate either accelerates or retards forgetting by speeding or slowing the turnover process. To date, the several experimental tests of this hypothesis are at variance with each other, and it is not yet clear whether brain metabolism is in anyway related to forgetting (Gleitman & Rozin, 1971). When this concept is considered in the developmental context, it first appears as though there might be a relationship between metabolic rate and forgetting, since it is well known that basal metabolic rate is much higher in childhood than adulthood (Keele & Neil, 1965). Brain metabolism, however, is not linked to body metabolism. In fact, it shows quite a different developmental pattern. Instead of decreasing during development, brain metabolism increases rapidly in the early postnatal stages to a maximum and then maintains that level into adulthood with only a slight decline as aging occurs. Thus the rapid rate of forgetting in infancy is not determined by a period of heightened brain metabolism.

The foregoing analysis of ontogenetic differences in memory can also be related to the decay theory of forgetting, now undergoing something of a resurgence in popularity (e.g., Gleitman, 1971; McGaugh & Dawson, 1971). That theory simply states that over time the memory trace, as represented in the central nervous system, diminishes in strength. In this context a major aim of Section III was to identify some of the central nervous system's mechanisms which might account for the rapid decay of memory in the neonate. Perhaps the identification of critical variables will be easier in those preparations where memory is declining most rapidly.

What is more important, however, is the fact that the effects many of these changes have on memory can now be investigated. For example, demyelinating agents (Bubis & Luse, 1964) and dietary deficiencies (Dobbing, 1964) can be used to alter concentration or rate of myelin deposition, central nervous system poisons such as 6 hydroxydopamine (Ungerstedt, 1968) can be used to destroy catecholaminergic synaptic clefts at different stages of development, various drugs can be used to block or enhance synaptic transmission, and higher brain centers can be temporarily paralyzed by topical application of potassium chloride.

Similarly, if more obviously, the effects of various environmental conditions on memory such as interpolated environmental complexity, body size relative to the environment, hormonally induced changes in species-specific behaviors, and many others are subject relatively easily to experimental study.

IV: RESOLUTION OF THE PARADOX

There are at least two possible resolutions to the paradoxical combination of poorer retention by immature mammals and the critical importance of early experience for adult behavior. One possibility, as noted in the introduction, is that key early experiences may permanently alter structure and function of the autonomic, endocrine, and central nervous systems.

Another possible resolution of this paradox is through mechanisms which act to maintain memories that are normally forgotten. The requirements of such a mechanism are that it be unavailable to the animal under controlled laboratory conditions, but at least potentially available to the animal in its natural environment. "Reinstatement" (Campbell & Jaynes, 1966) is such a mechanism. Campbell and Jaynes found that retention by immature rats was greatly facilitated if they were reexposed to the conditions of original training at points distributed throughout the retention interval. However, such reexposures by themselves were only a fractional portion of the experience necessary for original learning. Animals given only this "reexposure" training gave no evidence of learning.

The purpose in introducing the term "reinstatement" was to reemphasize the importance of such a phenomenon, rather than for innovation. In fact a similar mechanism for conserving memory was proposed long ago by Sechenov (1863). In discussing factors responsible for weakening of the memory trace, Sechenov noted the apparently poorer memory of immature humans:

> Indeed, what objects does the child remember? It remembers only those which are often before its eyes; the infant soon forgets even its mother should she die [p. 70].

Then he describes how a reinstatementlike mechanism might act to conserve retention in the adult:

> An adult, upon meeting another person even for a short time, receives from him a multitude of diverse sensations; the facial expression of the person, his features, pose, gait, manner of speaking, sound of voice, the subject of conversation, etc.—all remain in the memory for a more or less considerable time, depending on the strength of the impression; in the end, however, the traces of the impression begin to weaken. But one day another person is met, who along with other sensations induces one closely resembling the sensation produced by the first person. The latter sensation is revived in the memory, and becomes refreshed to a degree that it seems to be experienced anew. If these conditions are repeated from time to time, the traces remain [p. 70].

Sechenov then suggests that one reason for the poorer retention by children may result from the lesser opportunity for those experiences which conserve memory:

> In the child, however, these conditions, even if they are present, are incomparably less pronounced [p. 71].

Operationally, reinstatement is defined as better performance on a retention test by animals which have been briefly reexposed to the conditions of original learning during the retention interval in comparison with performance by animals receiving no such reexposure. Two critical assumptions underly the significance of this phenomenon. First, it is assumed that the reinstatement conditions are in themselves insufficient to establish learning. The second critical aspect of the reinstatement trials is their distribution throughout the retention interval—presumably if they were simply added as overtraining trials at the time of original learning, the benefit to retention would be a good deal less than is found with more even distribution throughout the retention interval. Both of these suppositions have been confirmed empirically, the first by Campbell and Jaynes (1966, 1969), Shubat and Whitehouse (1968), and Silvestri, Rohrbaugh, and Riccio (1970), and the second by Greenfield and Riccio (in press).

The generality of reinstatement has become increasingly apparent. Studies with the rat have established that reinstatement is effective in maintaining retention of "fear conditioning" (e.g., Silvestri et al., 1970), discrimination learning (Campbell & Jaynes, 1969), and in grade school children using a matching task (Hoving, Coates, Bertucci, & Riccio, 1972). In the latter experiment, children first learned to associate pairs of pictures, then eight weeks later they were asked to match pictures from a pool in accord with the original pairings. Children in the "reinstatement" condition were presented a story during the retention interval which included pairs of pictures matched as they had been during original learning. Eight weeks after learning, children given the reinstatement treatment required only about two-thirds as many trials to relearn the correct pairings as children not given this reinstatement treatment and about one-third as many trials as children given only the reinstatement treatment but not original learning.

Other variables which have been shown

to influence the magnitude of the reinstatement effect are the duration of each reinstatement trial and its content. Effects of these variables, though not uninteresting, are nevertheless straightforward and predictable. For example, Campbell and Jaynes (1969) trained weanling rats on a brightness discrimination, then measured their retention 10 weeks after original learning. Five groups of rats differed in their duration of exposure when reintroduced to this same task once each week during the retention interval: Either 0, 7.5, 15, 30, or 60 minutes of additional practice on the original brightness-discrimination task constituted the weekly reinstatement experience. Although 60 minutes of practice per week was nearly sufficient to eliminate the performance differences between rats given and those not given original learning, results of the other conditions permitted the conclusion of better retention the longer the duration of each reinstatement session.

Questions concerning which specific features of the reinstatement experience are critical for improving retention may also be relevant to the question, "What is forgotten?" Some cataloging of effective memory attributes probably is necessary for understanding animal memory (cf. Spear, 1971), and if certain elements of a conditioning trial are found to be both necessary and sufficient for reinstatement, these elements are in turn implicated as critical memory attributes. For example, Silvestri et al. (1970) found effective reinstatement of fear conditioning even though only the conditioned stimulus (without the unconditioned stimulus) was presented on each reinstatement trial. Silvestri et al. noted that the short duration of the conditioned-stimulus-only reinstatement trials was probably responsible for the occurrence of reinstatement instead of extinction, an assumption they then verified empirically. They also found that without the conditioned stimulus neither presentation of a "new" aversive stimulus nor the original unconditioned stimulus during reinstatement trials was sufficient to improve retention. Thus reinstatement of the conditioned stimulus dimension is probably more important than reinstatement of the unconditioned stimulus for maintenance of the target memory.

The fear-conditioning procedure used by Silvestri et al. involved alternate placement in the black and white compartments of a two-compartment box. The rat received footshocks in the black compartment but not in the white. Placement for an equivalent period into both compartments also constituted the reinstatement procedures which were effective even without the presence of shock in either compartment. Suppose the animals had been placed in only one of the compartments for reinstatement. Would reexposure to the presumably aversive compartment yield a different degree of reinstatement than reexposure to the previously more "safe" compartment? Greenfield (1969) conducted an appropriate experiment within this context and found reinstatement to be more effective when animals were placed only in the compartment in which they were previously shocked than if they were placed in the previously "safe" compartment. Of course, such a design does not permit identification of which memory attributes were forgotten independent of which were acquired originally. However, parallel results have been obtained in a reinstatement experiment by B. A. Campbell (in preparation) employing a brightness discrimination task. Campbell found reinstatement more effective when animals were exposed only to the nonreinforced stimulus during the retention interval than when animals were exposed only to the reinforced stimulus. Thus, in both experiments, the stimulus most effective in reinstatement was the stimulus associated with inhibition of responding. Whether this is due to more rapid forgetting of memory attributes associated with inhibition or to the fact that inhibition was more critical than excitation for original learning has yet to be determined.

V: IMPLICATIONS FOR THE ONTOGENY OF HUMAN MEMORY

The relevance of animal studies to human memory and cognitive processes has been subject to increasingly critical examination over the past decade. In the present instance, the same caution should be observed;

there are obviously vast differences in the qualitative aspects of behavioral development between rats and man. Man, for example, has an enormously greater capacity to encode and store complex information, and much of what is subsumed under the rubic "infantile amnesia" refers to the absence of memory for complex events and experiences that are easily recalled if learned in adulthood. Needless to say, we cannot equate the rapid forgetting of specific tasks such as passive-avoidance responding in rats with infantile amnesia in man. Moreover, there are substantial differences in rate of forgetting for different tasks in the rat. Some, like the passive-avoidance task, decline extremely rapidly, while others are forgotten much more slowly. Whether this reflects differences in strength of original learning, the relative susceptibility of inhibitory versus active responses to forgetting, or that some responses are more "natural" than others and therefore less likely to be forgotten is as yet unknown. But even if the basis for these differential rates of forgetting were known for the rat, it would still be an open question as to whether those same mechanisms operated in man.

On the other hand, there are a great many similarities between the ontogenesis of behavior and nervous system function in man and rat. Both are born with immature central nervous systems, are virtually helpless at birth, and are completely dependent on parental care. Similarly, both pass through periods of intense play and social interactions culminating in adult sexual behavior, and in both species complex social dominance hierarchies emerge during development.

Sequentially, the overall pattern of brain development is remarkably similar in rat and man. In both species, for example, the spinal motor roots are the first to myelinate and the cortical association areas the last. The time sequence is, of course, vastly different. At birth the human central nervous system is somewhat more advanced than that of the rat. In man, the spinal motor roots are fully myelinated at birth whereas myelinization does not begin in that area in the rat until two days postnatally. Similarly some brain stem and cerebellar structures show considerable myelinization at birth in man where no such development is evident in the rat. On the other hand, there is no sign of myelin in forebrain structures such as the thalamus, striatum, fornix, or cortex in either species. In man, the cortical association areas show traces of myelin at three months postnatally, but the rate of deposition is extremely slow, continuing past the tenth year postnatally and possibly into the third or fourth decade. Myelinization of the thalamic radiations begins about the same time but is completed by the eighth to ninth year. Other forebrain areas commence myelinization at about the same time, but deposition is much more rapid and nearly complete at 12 months of age (Yakovlev & Lecours, 1967). However by weight the amount of myelinization present in the frontal lobes is only 40% of the adult level at 10 months of age (O'Brien, 1970) indicating the extent to which myelin deposition continues in that area after the first year of life.

Other indexes of brain function confirm this general pattern of development. At birth, spontaneous EEG activity can be recorded in the human infant, although much of that activity is believed to be subcortical in origin (Ellingson, 1967). The pattern of EEG activity changes drastically during the first three postnatal years and then more gradually during the subsequent years into adolescence (Scheibel & Scheibel, 1964). Similarly, speed of nerve conduction continues to increase during development and fatigability decreases just as it does in the rat.

In summary, changes in central nervous system structure and function comparable to those occurring in the rat during the first months of life appear in man during the first years of life. For both species, these changes parallel the emergence of long-term memory as judged by improved retention of instrumental tasks in the rat and the disappearance of infantile amnesia in man. But whether central nervous system development is critical for the emergence of long-term memory is still an open question, since the sequence of behavioral development and the continuing acquisition of new responses could also account for the poor memory of infancy in both rat and man. The aim of this paper has been to review and analyze some of these changes in the hope it would stimulate research on the relative contribution of neurological and behavioral variables to the development of long-term memory. To the extent that future research on infrahuman animals points to either neurological or behavioral variables as critical for the emergence of long-term memory, the greater the necessity for examining those processes in man.

REFERENCES

Aghajanian, G. K., & Bloom, F. E. The formation of synaptic junctions in developing rat brain: A quantitative electron microscopic study. *Brain Research*, 1967, **6**, 716–727.

Agranoff, B. W. Recent studies on the stages of memory formation in the goldfish. In W. L. Byrne (Ed.), *Molecular approaches to learning and memory*. New York: Academic Press, 1970.

Agrawal, H. C., & Himwich, W. A. Amino acids, proteins and monoamines of the developing brain. In W. A. Himwich (Ed.), *Developmental neurobiology*. Springfield, Ill.: Charles C Thomas, 1970.

Altman, J. Postnatal neurogenesis and the problem of neural plasticity. In W. A. Himwich (Ed.), *Developmental neurobiology*. Springfield, Ill.: Charles C Thomas, 1970.

Allport, G. W. *Personality: A psychological interpretation*. New York: Holt, 1937.

Amster, H., Keppel, G., & Meyer, A. Learning and retention of letter pairs as a function of association strength. *American Journal of Psychology*, 1970, **83**, 22–39.

Bach, M. J., & Underwood, B. J. Developmental changes in memory attributes. *Journal of Educational Psychology*, 1970, **61**, 292–296.

Barondes, S. H. Some critical variables in studies of the effect of inhibitors of protein syntheses on memory. In W. L. Byrne (Ed.), *Molecular approaches to learning and memory*. New York: Academic Press, 1970.

Bass, N. H., Netsky, M. G., & Young, E. Microchemical studies of postnatal development in rat cerebrum: I. Migration and differentiation of cells. *Neurology*, 1969, **19**, 258–268.

Berger, B. D., & Stein, L. An analysis of learning deficits produced by scopolamine. *Psychopharmacologia*, 1969, **14**, 271–283. (a)

Berger, B. D., & Stein, L. Asymmetrical dissociation of learning between scopolamine and Wy 4036, a new benzodiazapine tranquilizer. *Psychopharmacologia*, 1969, **14**, 351–358. (b)

Bolles, R. C., & Woods, P. J. The ontogeny of behavior in the albino rat. *Animal Behavior*, 1964, **12**, 427–441.

Booth, D. A. Vertebrate brain ribonucleic acids and memory retention. *Psychological Bulletin*, 1967, **68**, 149–177.

Brazier, M. A. B. *A history of the electrical activity of the brain: The first half century*. New York: Macmillan, 1961.

Bubis, J. J., & Luse, S. A. An electron microscopic study of experimental encephalomyelites in the rat. *American Journal of Pathology*, 1964, **44**, 299–307.

Campbell, B. A. Developmental studies of learning and motivation in infraprimate mammals. In H. W. Stevenson, E. H. Hess, & H. L. Rheingold (Eds.), *Early behavior: Comparative and developmental approaches*. New York: Wiley, 1967.

Campbell, B. A., & Campbell, E. H. Retention and extinction of learned fear in infant and adult rats. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 1–8.

Campbell, B. A., & Jaynes, J. Reinstatement. *Psychological Review*, 1966, **73**, 478–480.

Campbell, B. A., & Jaynes, J. Effect of duration of reinstatement on retention of a visual discrimination learned in infancy. *Developmental Psychology*, 1969, **1**, 71–74.

Campbell, B. A., Jaynes, J. R., & Misanin, J. Retention of a light-dark discrimination in rats of different ages. *Journal of Comparative and Physiological Psychology*, 1968, **66**, 467–472.

Candland, D. K., & Campbell, B. A. Development of fear as measured by behavior in the open field. *Journal of Comparative and Physiological Psychology*, 1962, **5**, 593–596.

Crockett, W. H., & Noble, M. E. Age of learning, severity of negative reinforcement, and retention of learned responses. *Journal of Genetic Psychology*, 1963, **103**, 105–112.

Deese, J., & Hulse, S. H. *The psychology of learning*. (3rd ed.) New York: McGraw-Hill, 1967.

Deutsch, J. A., & Deutsch, D. *Physiological psychology*. Homewood, Ill.: Dorsey Press, 1966.

Deza, L., & Eidleberg, E. Development of cortical electrical activity in the rat. *Experimental Neurology*, 1967, **17**, 425–438.

Dobbing, J. The influence of early malnutrition on the development and myelination of the brain. *Proceedings of the Royal Society of Biology*, 1964, **159**, 503–509.

Dudycha, G. J., & Dudycha, M. M. Childhood memories: A review of the literature. *Psychological Bulletin*, 1941, **38**, 668–682.

Eayrs, J. T., & Goodhead, B. Postnatal development of the cerebral cortex in the rat. *Journal of Anatomy*, 1959, **93**, 385–401.

Ellingson, R. J. The study of brain electrical activity of infants. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in child development and behavior*. New York, Academic Press, 1967.

Feigley, D. A., & Spear, N. E. Effect of age and punishment condition on long-term retention by the rat of active- and passive-avoidance learning. *Journal of Comparative and Physiological Psychology*, 1970, **73**, 515–526.

Flechsig, P. *Gehirn & Seele*, Leipzeig, 1896.

Flexner, L. B., Flexner, J. B., & Roberts, R. B.

Memory in mice analyzed with antibiotics. *Science,* 1967, **155**, 1377–1383.

Friede, R. L. *Topographic brain chemistry.* New York: Academic Press, 1966.

Frieman, J. P., Rohrbaugh, M., & Riccio, D. C. Age differences in the control of acquired fear by tone. *Canadian Journal of Psychology,* 1969, **23**, 237–244.

Gleitman, H. Forgetting of long-term memories in animals. In W. K. Honig & P. H. R. James (Eds.), *Animal memory.* New York: Academic Press, 1971.

Gleitman, H., & Rozin, P. Learning and memory. *Fish Physiology,* 1971, **6**, 196–278.

Goulet, L. R. Verbal learning in children: Implications for developmental research. *Psychological Bulletin,* 1968, **69**, 359–376.

Grafstein, B. Postnatal development of the corpus collosum in the cat. In P. Kellaway & I. Peterson (Eds.), *Neurological and electroencephalographic correlative studies in infancy.* New York: Grune & Stratton, 1964.

Greenfield, H., & Riccio, D. C. Conditioned reinstatement in rats: Effect of exposure distribution and cue. *Psychological Reports,* 1972, in press.

Harlow, H. F. The heterosexual system in monkeys. *American Psychologist,* 1962, **17**, 1–9.

Harlow, H. F., & Harlow, M. K. Social deprivation in monkeys. *Scientific American,* 1962, **207**, 136–146.

Himwich, W. A. *Developmental neurobiology.* Springfield, Ill.: Charles C Thomas, 1970.

Hinde, R. A. *Bird vocalizations in relation to current problems in biology and psychology.* New York: Cambridge University Press, 1969.

Hoving, K. L., Coates, L., Bertucci, M., & Riccio, D. C. Reinstatement of effects in children. *Developmental Psychology.* 1972, in press.

Hutchinson, R. R., Ulrich, R. E., & Azrin, N. H. Effects of age and related factors on the pain-aggression reaction. *Journal of Comparative and Physiological Psychology,* 1965, **59**, 365–369.

Inhelder, B. Memory and intelligence in the child. In D. Elkind & J. H. Flavell (Eds.), *Studies in cognitive development.* Oxford: Oxford University Press, 1969.

Jacobson, S. Sequence of myelinization in the brain of the albino rat. A. Cerebral cortex, thalamus and related structures. *Journal of Comparative Neurology,* 1963, **121**, 5–29.

Keele, C. A., & Neil, E. *Samson Wright's applied physiology.* (11th ed.) Oxford: Oxford University Press, 1965.

Kendler, H. H., & Kendler, T. S. Vertical and horizontal processes in problem solving. *Psychological Review,* 1962, **69**, 1–16.

Keppel, G. Verbal learning in children. *Psychological Bulletin,* 1964, **61**, 63–80.

Kirby, R. H. Acquisition, extinction, and retention of an avoidance response in rats as a function of age. *Journal of Comparative and Physiological Psychology,* 1963, **56**, 158–162.

Klein, S. B., & Spear, N. E. Influence of age in short-term retention of active avoidance learning in rats. *Journal of Comparative and Physiological Psychology,* 1969, **69**, 583–589.

Levine, S., & Mullins, R. F., Jr. Hormones in infancy. In G. Newton & S. Levine (Eds.), *Early experience and behavior.* Springfield, Ill.: Charles C Thomas, 1968.

Levy, D. M. The infant's earliest memory of inoculation: A contribution to public health procedures. *Journal of Genetic Psychology,* 1960, **96**, 3–46.

Mandel, P., Rein, H., Harth-Edel, S., & Mardell, R. Distribution and metabolism of ribonucleic acid in the vertebrate central nervous system. In D. Richter (Ed.), *Comparative neurochemistry.* New York: Pergamon Press 1964.

McGaugh, J. L., & Dawson, R. G. Modification of memory storage processes. In W. K. Honig & P. H. R. James (Eds.), *Animal memory.* New York: Academic Press, 1971.

McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* (2nd ed.) New York: Longmans, Green, 1952.

Moorcroft, W. H., Lytle, L. D., & Campbell, B. A. Ontogeny of starvation-induced behavioral arousal in the rat. *Journal of Comparative and Physiological Psychology,* 1971, **75**, 59–67.

O'Brien, J. S. Lipids and myelination. In W. A. Himwich (Ed.), *Developmental neurobiology.* Springfield, Ill.: Charles C Thomas, 1970.

Paré, W. P. Age, sex, and strain differences in the aversive threshold to grid shock in the rat. *Journal of Comparative and Physiological Psychology,* 1969, **69**, 214–218.

Paulson, G., & Gottlieb, G. Development reflexes: The reappearance of foetal and neonatal reflexes in aged patients. *Brain Research,* 1968, **91**, 37–52.

Perkins, C. C., Jr. A conceptual scheme for studies for stimulus generalization. In D. I. Mostofsky (Ed.), *Stimulus generalization.* Stanford: Stanford University Press, 1965.

Peterson, N. Effect of monochromatic rearing on the control of responding by wavelength. *Science,* 1962, **136**, 774–775.

Postman, L. Transfer, interference and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlossberg's experimental psychology.* (3rd ed.) New York: Holt, Rinehart & Winston, 1971.

Prokasy, W. F., & Hall, J. F. Primary stimulus generalization. *Psychological Review,* 1963, **70**, 310–322.

Purpura, D. P., Shofer, R. J., Housepian, E. M., & Noback, C. R. Comparative ontogenesis of structure-formation relationships in cerebral and cerebellar cortex. In D. P. Purpura & J. P. Schade (Eds.), *Progress in brain research.* Vol. 4. *Growth and maturation of the brain.* Amsterdam: Elsevier, 1964.

Ratner, S. C. Annelids and learning: A critical review. In W. C. Corning & S. C. Ratner (Eds.), *Chemistry of learning.* New York: Plenum Press, 1967.

Riesen, A. H. Stimulation as a requirement for growth and function in behavioral development. In D. W. Fiske & S. R. Maddi, (Eds.), *Functions of varied experience.* Homewood, Ill.: Dorsey Press, 1961.

ROBINSON, N. Histochemistry of rat brainstem monamine oxidase during maturation. *Journal of Neurochemistry*, 1968, **15**, 1151–1158.

ROSENZWEIG, M. R., BENNETT. E. L., DIAMOND, M. C., WU, S., SLAGLE, R. W., & SAFFRAM, E. Influence of environmental complexity and visual stimulation on development of occipital cortex in rats. *Brain Research*, 1969, **14**, 427–445.

ROZENSWEIG, M. R., KRECH, D., BENNETT, E. L., & DIAMOND, M. C. Modifying brain chemistry and anatomy by enrichment or impoverishment of experience. In G. Newton & S. Levine (Eds.), *Early experience and behavior.* Springfield, Ill.: Charles C Thomas, 1968.

RUDOLF, R. L., HONIG, W. K., & GERRY, J. E. Effects of monochromatic rearing on the acquisition of stimulus control. *Journal of Comparative and Physiological Psychology*, 1969, **67**, 50–57.

SAUL, L. J., SNYDER, T. R., & SHEPPARD, E. On earliest memories. *Psychoanalytic Quarterly*, 1956, **25**, 228–237.

SCHACHTEL, E. G. On memory and childhood amnesia. *Psychiatry*, 1947, **10**, 1–26.

SCHEIBEL, M. E., & SCHEIBEL, A. B. Some neural substrates of postnatal development. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research.* New York: Russell Sage Foundation, 1964.

SCHERRER, J. Electrophysiological aspects of cortical development. In E. A. Asraytan (Ed.), *Progress in brain research.* Amsterdam: Elsevier, 1967.

SCHULENBURG, C. J., RICCIO, D. C., & STIKES, E. R. Acquisition and retention of a passive-avoidance response as a function of age in rats. *Journal of Comparative and Physiological Psychology*, 1971, **74**, 75–83.

SECHENOV, I. M. *Reflexes of the brain.* Cambridge: M.I.T. Press, 1965. (Originally published, 1863.)

SEIDEN, L. S., & PETERSON, D. D. Blockage of L-dopa reversal of reserpine-induced conditioned avoidance response suppression by disulfiram. *Journal of Pharmacology and Experimental Therapeutics*, 1968, **163**, 84–90.

SHUBAT, E. E., & WHITEHOUSE, J. M. Reinstatement: An attempt at replication. *Psychonomic Science*, 1968, **12**, 215–216.

SILVESTRI, R., ROHRBAUGH, M. J., & RICCIO, D. C. Conditions influencing the retention of learned fear in young rats. *Developmental Psychology*, 1970, **2**, 389–395.

SLUCKIN, W. *Imprinting and early learning.* Chicago: Aldine, 1965.

SMITH, N. Effects of interpolated learning on the retention of an escape response in rats as a function of age. *Journal of Comparative and Physiological Psychology*, 1968, **65**, 422–426.

SPEAR, N. E. Verbal learning and retention. In M. R. D'Amato, *Experimental psychology: Psychophysics, methodology and learning.* New York: McGraw-Hill, 1970.

SPEAR, N. E. Forgetting as retrieval failure. In W. K. Honig & H. James (Eds.), *Animal memory.* New York: Academic Press, 1971.

STEIN, L., & BERGER, B. D. Paradoxical fear-increasing effects of tranquilizers: Evidence of repression of memory in the rat. *Science*, 1969, **166**, 253–256.

THOMPSON, R. W., KOENIGSBERG, L. A., & TENNISON, J. C. Effects of age on learning and retention of an avoidance response in rats. *Journal of Comparative and Physiological Psychology*, 1965, **60**, 457–459.

TOMKINS, S. S. *Affect, imagery, consciousness.* Vol. 2. *The negative affects.* New York: Springer, 1963.

UNDERWOOD, B. J. Interference and forgetting. *Psychological Review*, 1957, **64**, 49–60.

UNDERWOOD, B. J. Degree of learning and the measurement of forgetting. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 112–129.

UNDERWOOD, B. J. *Experimental psychology.* (2nd ed.) New York: Appleton-Century-Crofts, 1966.

UNDERWOOD, B. J. Attributes of memory. *Psychological Review*, 1969, **76**, 559–573.

UNDERWOOD, B. J., & EKSTRAND, B. R. An analysis of some shortcomings in the interference in forgetting. *Psychological Review*, 1966, **73**, 540–549.

UNDERWOOD, B. J., & FREUND, J. S. Effect of temporal separation of two tasks on proactive inhibition. *Journal of Experimental Psychology*, 1968, **78**, 50–54.

UNDERWOOD, B. J., & POSTMAN, L. Extraexperimental sources of interference in forgetting. *Psychological Review*, 1960, **60**, 457–459.

UNGERSTEDT, U. 6-Hydroxydopamine induced degeneration of central monoamine neurons. *European Journal of Pharmacology*, 1968, **5**, 107–110.

WALDFOGEL, S. The frequency and affective character of childhood memories. *Psychological Monographs*, 1948, **62**(4, Whole No. 291).

WALEN, S. R. Recall in children and adults. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 94–98.

WIESEL, T. N., & HUBEL, D. H. Comparison of the effects of unilateral and bilateral eye closure on cortical unit response in kittens. *Journal of Physiology*, 1965, **28**, 1029–1041.

YAKOVLEV, P. I., & LECOURS, A. The myelogenetic cycles of regional maturation of the brain. In A. Minkowski (Ed.), *Regional development of the brain in early life.* Philadelphia: Davis, 1967.

YERKES, R. M. The intelligence of earthworms. *Journal of Animal Behavior*, 1912, **2**, 332–352.

(Received July 21, 1971; revision received December 13, 1971)

Psychological Review
1972, Vol. 79, No. 3, 237–244

MECHANISMS OF HABITUATION IN THE BRAIN STEM[1]

PHILIP M. GROVES[2]
University of Colorado

GARY S. LYNCH
Department of Psychobiology, University of California, Irvine

The brain stem reticular formation is proposed to underlie short-term response habituation and sensitization in intact organisms. The organization of the reticular formation is viewed, in light of both physiological and anatomical observations, in terms of an outer layer of small cells, the parvicellular region, surrounding an inner core of giant cells. Sensory information can be transmitted to the parvicellular associational neurons, directly or via cortical relays, where according to the model, it becomes subject to the synaptic processes associated with habituation and sensitization. The modified output is then channeled via the giant-celled reticular core to the various final common paths. The implications of this theoretical view are discussed.

The brain stem reticular formation has played a prominent role in the neural systems analysis of short-term response habituation.[3] Its specific role in response decrement, however, has varied considerably among theories. Sokolov (1963) hypothesized that the reticular formation served an elaborative function in generating the orienting response and that habituation of this response was due to corticofugal blockade of input into the reticular formation. Hernández-Peón (1960), on the other hand, suggested that with repetitive sensory stimulation, the reticular formation can block afferent input as far peripherally as the first sensory relay nucleus. Sharpless and Jasper (1956) proposed that the reticular formation itself was the site of the decremental process for the arousal response, and later Stein (1966) hypothesized that the process might be recurrent inhibition of excitatory reticular activation. Thus, the involvement of the reticular formation in explanations of response decrement to repeated sensory stimulation has been a persistent theme for many years.

Neurophysiological evidence tends to support this theoretical emphasis on the role of the reticular formation in habituation. Habituation of cellular activity has been reported to occur reliably in the area of the brain stem reticular formation, particularly in the mesencephalon (Bell, Sierra, Buendia, & Segundo, 1964; Groves, Miller, & Parker, 1972; Hirano, Best, & Olds, 1970; Horn & Hill, 1966; Scheibel & Scheibel, 1965) even in the absence of changes in primary evoked potentials (Weinberger, Goodman, & Kitzes, 1969). In addition, studies on the effects of stimulation or ablation of mesencephalic reticular formation on habituation support the hypothesis that it is involved in response habituation (Capps & Stockwell, 1968; Glickman & Feldman, 1961). In contrast, decrement to repetitive stimulation in a *primary* sensory system (e.g., Kitzes & Buchwald, 1969; Webster, 1969) does not appear to be necessarily related to the more rapid time course of behavioral response habituation (see Groves & Thompson, 1970; Thompson & Spencer, 1966; Weinberger et al., 1969).

---

[1] This work was supported in part by Grant NA 07661 from the National Institute of Neurological Diseases and Stroke and Grant MH 11279 from the National Institute of Mental Health awarded to Richard F. Thompson; Grant MH NS 19515-01 from the National Institute of Mental Health to Philip M. Groves; and Grant BO 16973 from the National Science Foundation to Gary S. Lynch. We wish to thank Richard F. Thompson for his many suggestions and criticisms, and Seth Sharpless, Stephen Glickman, and Norman Weinberger for their helpful comments on an earlier version of this article.

[2] Requests for reprints should be sent to Philip M. Groves, Department of Psychology, University of Colorado, Boulder, Colorado 80302.

[3] There is now mounting indirect evidence which suggests that habituation retained across days (i.e., the form of decrement commonly referred to as "between-sessions" habituation, which may take days or weeks to recover) involves elaboration by forebrain structures, most particularly frontal (or other association) cortex and hippocampus, but possibly other forebrain structures as well. Thus, ablation of these areas or functional blockade of their effects abolishes or significantly retards between-sessions habituation (e.g., Roydes, 1970; Thompson & Welker, 1963). It cannot be ascertained at present whether these effects are secondary to the effects on within-sessions habituation which we presume is mediated by reticular formation. However, studies of the effects of interstimulus interval on within- versus between-sessions habituation provide supporting evidence that these two different time-course phenomena are somewhat independent. Within-sessions habituation is directly related to rate of stimulus presentation, while between-sessions habituation (i.e., across days) is inversely related to rate of stimulus presentation (e.g., Carew, Pinsker, & Kandel, 1971; Davis, 1970a, 1970b). We use the term "habituation" here to refer only to the process of habituation which occurs usually within minutes rather than days, which is the time course of habituation most commonly studied.

In spite of this evidence implicating the reticular formation in response habituation, there has been little discussion of how various cellular elements, which comprise the reticular network, are involved. This lack is critical in view of recent anatomical and physiological studies showing that the reticular formation contains many different types of neurons (Brodal, 1957; Ramón-Moliner & Nauta, 1966; Scheibel & Scheibel, 1967; Valverde, 1961, 1962) and has subdivisions which serve different, and in some cases opposing, functions (Brodal, 1957; Sauerland, Nakamura, & Clemente, 1967). In this paper, we propose a specific model of how the reticular formation may be involved in response habituation. Analogous structural elements of the reticular formation and the spinal cord reflect a number of functional similarities including mediation of response plasticity to repetitive sensory stimuli.

Habituation in the Spinal Animal

It has been known for some time that habituation could be mediated entirely at the spinal level (Prosser & Hunter, 1936; Sherrington, 1906; Thompson & Spencer, 1966). Thompson and Spencer (1966), on the basis of an extensive series of experiments on habituation of the flexion reflex of acute spinal cat (Spencer, Thompson, & Nielson, 1966a, 1966b, 1966c), proposed that habituation of the flexion reflex was due to the balance of two processes, one excitatory (sensitization) and the other decremental (habituation). They suggested that plasticity in the spinal animal was based on the action of interneurons, and subsequent studies verified these predictions (Groves, DeMarco, & Thompson, 1969; Groves, Glanzman, Patterson, & Thompson, 1970; Groves & Thompson, 1970; Wickelgren, 1967). Thus, habituation and sensitization appear to be due to the actions of two relatively distinct populations of interneurons in the dorsal grey matter of the spinal cord. One population shows only progressive decrement in response to repeated stimulation (Type H) while the other population shows increment followed by variable decrement (Type S). A dual-process view of plasticity to repeated stimulation was thus developed, and it was suggested that similar classes of interneurons in the brain might mediate habituation and sensitization in the intact organism (Groves & Thompson, 1970).

Comparative Organization of the Reticular Formation and Spinal Cord

It has been often noted that the phylogenetically "primitive" reticular formation (Scheibel & Scheibel, 1967) represents a continuation of the anatomical arrangements found in the spinal cord and includes many of the neuronal cell types found there (Matsushita, 1969; Ramón-Moliner & Nauta, 1966). Brodal (1957), utilizing Nissl cell staining methods, pointed out that the reticular formation may be subdivided into a "giant-celled" core surrounded in part by a blanket or layer of smaller cells (the parvicellular region). These two divisions can also be distinguished according to their dentritic and axonal orientations. The smaller-celled regions have dendrites directed toward neighboring sensory systems and axonal projections into the giant-celled core; Brodal hypothesized that these parvicellular elements might serve an associational function. The giant cells of the core have dendritic proliferations of great size which seem to ramify in an almost random fashion, although their rostral-caudal dimensions are possibly restricted (Scheibel & Scheibel, 1967). The axons of these cells travel great distances emitting collaterals all along their course. Brodal hypothesized that the giant cells served as effectors for the reticular network. Golgi studies have, in general, verified these descriptions (Scheibel & Scheibel, 1958; Valverde, 1961).

The parallels between this arrangement and that of the spinal cord are striking. Thus, Valverde (1961) suggests that "the lateral reticular areas seem to play in the brain stem the same role as the zona intermedia and basal parts of the dorsal horn in the spinal cord [p. 47]." Hence, the spinal cord has layers of small- and medium-sized cells dorsally placed in the grey matter that appear in part to relay cutaneous input ultimately to the ventrally placed large motoneurons. The dorsally placed cells appear to subserve the "associational" and behavioral plasticity functions, while the more ventrally placed interneurons and the montoneurons serve as points of interaction and output for these behavioral processes (Groves & Thompson, 1970). Similarly, in the reticular formation, a dorso laterally placed region of interneurons (parvicellularis) receives various inputs from neighboring sensory systems and relays information to the giant cells, the final common output neurons. A more specific analogy can be seen in the course of the fifth cranial nerve. Cutaneous input from the facial regions travels primarily in the descending root of the fifth cranial nerve to terminate in the spinal nucleus of the fifth nerve and/or the nearby nucleus reticularis parvicellularis (Darian-Smith, 1966). From these associational stations, short axons arise which project onto cells of the more centrally located medullary and pontine reticular formation. Giant cell axons then innervate motor systems through the neuraxis. Our hypothesis is that elements analogous to the Type H and Type S interneurons in spinal cord are located in these parvicellular regions. The output function might then be the result of the action of these labile parvicellular elements on the giant cells—effector neurons for distribution to various final common pathways throughout the neuraxis.

It is particulary important at this point to consider several specific aspects of the anatomical arrangement described above and how they might relate to habituation phenomena. Habituation occurs in all modalities (with a possible exception discussed below), yet it is not certain if all sensory systems utilize parvicellular relays in reaching the reticular formation core. Partly this stems from the fact that the boundaries of what is usually called reticular formation are blurred. This is particularly evident, for example, in the case of the junction of the nucleus reticularis parvicellularis and the spinal nucleus of the fifth cranial nerve. Furthermore, the relationship between the small-celled and medium-sized-celled reticular areas in the mesencephalon and thalamus and the giant-celled reticular core has not received detailed study. However, recent work demonstrates that the small-celled regions very probably innervate the more caudal giant-celled reticular core (Scheibel & Scheibel, 1967; Schlag & Challet, 1963; Waldron & Gwyn, 1969; Weinberger, Velasco, & Lindsley, 1965). Although these anterior regions are not typically thought of as associational cells for the giant-celled reticular core, it is quite possible from both anatomical and neurophysiological evidence that they serve this function. If so, they would vastly increase the extent to which sensory input could make mediated or "associational" contact with the reticular core. Finally, recent studies (Mizumo, Clemente, & Sauerland, 1968; Valverde, 1962) have demonstrated that some cortical sensory regions innervate the smaller-celled reticular areas of the brain stem (including mesencephalic areas) thus opening the possibility that some sensory input to the parvicellular regions may involve a cortical relay. This could have the important functional significance of allowing for certain cortical influences on habituation,

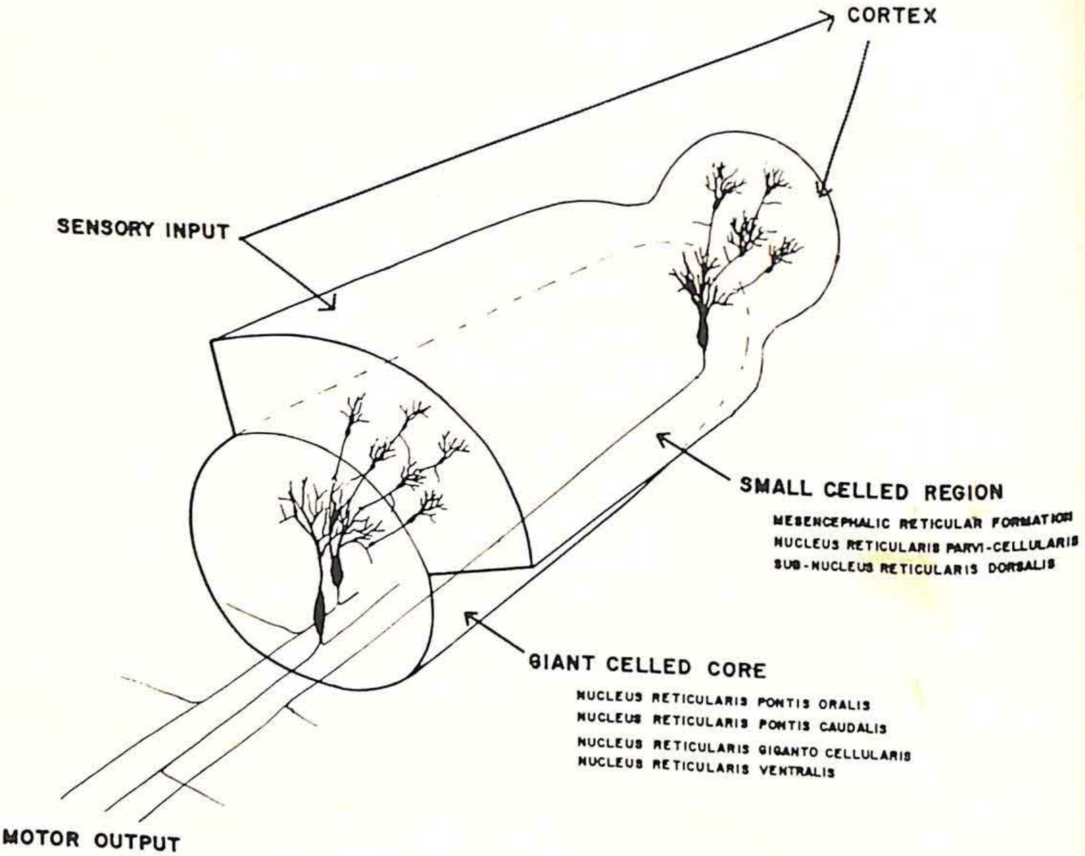

CORTEX
SENSORY INPUT
SMALL CELLED REGION
MESENCEPHALIC RETICULAR FORMATION
NUCLEUS RETICULARIS PARVI-CELLULARIS
SUB-NUCLEUS RETICULARIS DORSALIS
GIANT CELLED CORE
NUCLEUS RETICULARIS PONTIS ORALIS
NUCLEUS RETICULARIS PONTIS CAUDALIS
NUCLEUS RETICULARIS GIGANTO CELLULARIS
NUCLEUS RETICULARIS VENTRALIS
MOTOR OUTPUT

FIG. 1. Summary diagram of the reticular formation as conceptualized in the role of underlying response habituation. (The giant-celled core represents an output system for the reticular formation. The small-celled region contains the associational elements which receive sensory input [dotted lines] from either subcortical or cortical stations. These elements are proposed as the neuronal substrates of habituation, their influences being relayed throughout the neuraxis by means of their connections with the giant-celled core. Various specific nuclei contained within these two major functional subdivisions of the reticular formation are listed.)

for example, those aspects of habituation involving sensory discriminations (Sharpless & Jasper, 1956).

In addition to the various relayed sensory effects, the giant-celled reticular core receives direct sensory innervation. The most notable example of this is provided by the small diameter ascending spinal fibers carried in the ventrolateral funiculus. A considerable body of anatomical (Bowsher, Mallart, Petit, & Albe-Fessard, 1968) and physiological (Casey, 1969; Melzack & Wall, 1965) data link the spinoreticular fibers with the mediation of pain. In view of experiments showing that little or no habituation occurs to painful stimuli (Sokolov, 1963), it may be that this system terminating directly on the reticular core does not habituate.

Summary of the Model

Figure 1 summarizes in schematic fashion the neuroanatomical and functional relations we have described. For convenience, we have not drawn the direct sensory input to the giant-celled reticular region. What has typically been dealt with as reticular formation, namely, the mesencephalic reticular formation, is pictured as an anterior functional extension of the parvicellular associational system. Sensory information can be transmitted through the parvicellular associational neurons, directly or via cortical relays, where it becomes subject to the synaptic processes associated with sensitization and habituation. This modification is then channeled via the giant-celled reticular core to the various final common pathways. It is important to indicate that there are in all likelihood parallel inputs and outputs to elements within the reticular formation and many of these may not subserve a plasticity function, as many elements do not at the spinal level. As far as synaptic mechanisms are concerned, present evidence from a variety of model system analyses suggests that active synaptic inhibition is not involved but rather that habituation may result from some form of synaptic depression and that sensitization may result from incremental synaptic activity (see Groves & Thompson; Thompson & Spencer, 1966).

Some Reinterpretations

While the view that the reticular formation is involved in short-term plasticity of response to repetitive sensory stimulation is by no means a novel one, our model does provide a clear anatomical substrate for many well-known effects of iterated sensory stimulation. One would expect, for example, to observe consistent examples of plasticity of response to repetitive stimulation in reticular formation neurons. Habituation and sensitization occur reliably in cells of the reticular formation with a time course similar to these phenomena in the intact organism (Bell et al., 1964; Groves et al., 1972; Hirano et al., 1970; Horn & Hill, 1966; Scheibel & Scheibel, 1965; Weinberger et al., 1969). Additionally, interactions of modalities on the processes of habituation and sensitization are easily understood in our view of the functional organization of the reticular core of the brain. Hence, while habituation may occur to one modality, response strength can be transiently restored by a strong, extra stimulus of the same or a different modality. This convergence of effects and its distribution to various levels of the neuraxis are easily understood on the basis of the convergence of effects onto, and distribution by, the large-celled reticular core. The clear interaction of arousal or "state" with habituation of behavioral and neural responses (see Groves & Thompson, 1970) is also easily and obviously accounted for on the basis of a reticular formation origin for these processes. In addition to the clear occurrence of habituation and sensitization processes and their interaction in the reticular formation, the parametric characteristics of response habituation (see Thompson & Spencer, 1966) have been shown to hold for reticular neurons (Bell et al., 1964; Groves et al., 1972; Scheibel & Scheibel, 1965).

The effects of sleep on habituation (see Sokolov, 1963; Johnson & Lubin, 1967) also have a clear anatomical substrate in our model of reticular formation function. It has been known for some time that many habituated responses show a spontaneous recovery with the onset of sleep. Groves and Thompson (1970) suggested that this reflected a reduction in function of systems primarily involved in response plasticity in the wake organism. If this system were the reticular formation, then one would expect little or no habituation of reticular units during sleep. Scheibel and Scheibel (1965) have shown that repetitive activation of the sciatic nerve, which normally results in pronounced decrement in evoked discharges of single cells in reticular formation of unanesthetized cat, produces little or no decrement during slow-wave sleep. While the mechanisms of this alteration in lability of reticular neurons are as yet unknown, the fact of its occurrence in view of a reticular substrate for short-term behavioral plasticity is indeed suggestive.

Another puzzling phenomenon for many neural systems theories of habituation is the "missing-stimulus effect" described by Sokolov (1963). For some responses, the *absence* of a stimulus may produce an increase in response similar to alteration of any other characteristic of the stimulus. Groves and Thompson (1970) suggested that this was a phenomenon characteristic of state variables and could most easily be accounted for on the basis of temporal conditioning of state. Interestingly, it was pointed out that temporal conditioning may be limited to responses that clearly reflect level of arousal. In addition to this, single cells in reticular formation have been observed to respond to a regularly repeated stimulus after it has been terminated at intervals which closely approximate the training intervals (Scheibel & Scheibel, 1965). These observations, together with the theoretical view of Groves and Thompson, are most easily explained by assuming a reticular origin for the missing-stimulus effect.

Evidence derived from the effects of stimulation and ablation of mesencephalic reticular formation (our rostral extension of the labile, parvicellular elements of the reticular formation) is entirely consistent with its proposed role in response plasticity. Ablation of mesencephalic reticular formation has been shown to abolish short-term habituation of the startle response in rats apparently without affecting response amplitude (Capps & Stockwell, 1968). Further, habituation of electrencephalographic arousal to direct electrical stimulation of the mesencephalic formation of rats has been demonstrated (Glickman & Feldman, 1961). Direct electrical stimulation in the region of the nucleus gigantocellularis, on the other hand, elicits escape behavior in cats (Casey, 1971) and rats (Keene & Casey, 1970).

Hippocampectomy, amygdalectomy, decortication, frontal cortical ablation, and other partial neocortical ablations have been reported to retard response habituation (see Thompson & Spencer, 1966). Within the present theoretical framework, these effects can be viewed as secondary to changes in reticular excitability which many of these manipulations are known to produce. Many forebrain structures which appear to influence habituation (e.g., hippocampus and frontal cortex) have intimate anatomical connections with the reticular formation (see Lynch, 1970; Valverde, 1962), and there is considerable evidence that these suppress neurophysiological (see reviews by Dell, 1963; Sterman & Clemente, 1967) and behavioral signs of reticular arousal (Lynch, Ballantine, & Campbell, 1969). It is reasonable, therefore, that for the most part, these ablations retard, but do not abolish, short-term habituation of many behavioral responses.

Although theories of response habituation have been proposed which deal very effectively with many of the characteristics of habituation, a clear and adequate anatomical substrate for short-term plasticity in the intact organism has yet to be proposed. The possibility that the reticular formation, an area modulated by a large number of neurological systems, is central to habituation could explain the wide range of conditions known to influence these processes. More important, perhaps, than mediation of simple forms of plasticity by the brain stem reticular formation is the possibility that parallel processing of these forms of plasticity at all lower levels of the neuraxis may involve interneuronal mechanisms which are qualitatively similar through the continuous central core of the spinal cord and the brain.

REFERENCES

Bell, C., Sierra, G., Buendia, N., & Segundo, J. P. Sensory properties of neurons in the mesencephalic reticular formation. *Journal of Neurophysiology*, 1964, **27**, 961–987.

Bowsher, D., Mallart, A., Petit, D., & Albe-Fessard, D. A bulbar relay to the centre median. *Journal of Neurophysiology*, 1968, **31**, 288–300.

Brodal, A. *The reticular formation of the brain stem. Anatomical aspects and functional correlations.* Edinburg: Oliver & Boyd, 1957.

Capps, M. J., & Stockwell, C. W. Lesions in the midbrain reticular formation and the startle response in rats. *Physiology and Behavior*, 1968, **3**, 661–665.

Carew, T., Pinsker, H., & Kandel, E. Long-term habituation of siphon withdrawal reflex in *Aplysia*. Paper presented at the XXV International Congress of Physiological Sciences, Munich, August 1971.

Casey, K. Somatic stimuli, spinal pathways, and size of cutaneous fibers influencing unit activity in the medial medullary reticular formation. *Experimental Neurology*, 1969, **25**, 35–56.

Casey, K. L. Somatosensory responses of bulboreticular units in awake cat: Relation to escape-producing stimuli. *Science*, 1971, **173**, 77–80.

Darian-Smith, I. Neural mechanisms of facial sensation. *International Review of Neurobiology*, 1966, **9**, 301–395.

Davis, M. Effects of interstimulus interval length and variability on startle-response habituation in the rat. *Journal of Comparative* *and Physiological Psychology*, 1970, **72**, 117–192. (a)

Davis, M. Interstimulus interval and startle response habituation with a "control" for total time during training. *Psychonomic Science*, 1970, **20**, 39–41. (b)

Dell, P. Reticular homeostasis and cortical reactivity. In G. Moruzzi, A. Fessard, & H. Jasper (Eds.), *Progress in brain research.* Vol 1. *Brain mechanisms.* New York: Elsevier, 1963.

Glickman, S. E., & Feldman, S. M. Habituation of the arousal response to direct stimulation of the brain stem. *Electroencephalography and Clinical Neurophysiology*, 1961, **13**, 703–709.

Groves, P. M., DeMarco, R., & Thompson, R. F. Habituation and sensitization of spinal interneuron activity in acute spinal cat. *Brain Research*, 1969, **14**, 521–525.

Groves, P. M., Glanzman, D. L., Patterson, M. M., & Thompson, R. F. Excitability of cutaneous afferent terminals during habituation and sensitization in acute spinal cat. *Brain Research*, 1970, **18**, 388–392.

Groves, P., Miller, S. W., & Parker, M.V. Habituation and sensitization of interneuron activity in reticular formation of rat. *Physiology and Behavior*, 1972, in press.

Groves, P. M., & Thompson, R. F. Habituation: A dual-process theory. *Psychological Review*, 1970, **77**, 419–450.

Hernández-Péon, R. Neurophysiological correlates of habituation and other manifestations of plastic inhibition (internal inhibition). In H. H. Jasper & G. D. Smirnov (Eds.), The Moscow colloquium on electroencephalography of higher nervous activity. *Electroencephalography and Clinical Neurophysiology*, 1960, Suppl. 13.

Hirano, T., Best, P., & Olds, J. Units during habituation, discrimination learning and extinction. *Electroencephalography and Clinical Neurophysiology*, 1970, **28**, 127–135.

Horn, G., & Hill, R. M. Responsiveness to sensory stimulation of units in the superior colliculus and subjacent tectotegmental regions in the rabbit. *Experimental Neurology*, 1966, **14**, 199–223.

Johnson, L. C., & Lubin, A. The orienting reflex during waking and sleeping. *Conditional Reflex*, 1967, **2**, 160.

Keene, J. J., & Casey, K. L. Excitatory connection from lateral hypothalamic self-stimulation sites to escape sites in medullary reticular formation. *Experimental Neurology*, 1970, **28**, 155.

Kitzes, M., & Buchwald, J. Progressive alternation in cochlear nucleus, inferior colliculus, and medial geniculate body responses during acoustic habituation. *Experimental Neurology*, 1969, **25**, 85–105.

Lynch, G. Separable forebrain mechanisms controlling different manifestations of spontaneous activity. *Journal of Comparative and Physiological Psychology*, 1970, **70**, 48–59.

Lynch, G., Ballantine, P., & Campbell, B. A. Potentiation of behavioral arousal following cortical lesions and subsequent recovery. *Experimental Neurology*, 1969, **23**, 195–206.

Matsushita, M. Some aspects of the interneuronal connections in cats spinal gray matter. *Journal of Comparative Neurology*, 1969, **136**, 57–80.

Melzack, R., & Wall, P. D. Pain mechanisms: A new theory. *Science*, 1965, **150**, 971–974.

Mizumo, N., Clemente, C., & Sauerland, E. Projections from the orbital gyrus in the cat, to brain stem structures. *Journal of Comparative Neurology*, 1968, **133**, 463–476.

Prosser, C. L., & Hunter, W. S. The extinction of startle responses and spinal reflexes in the white rat. *American Journal of Physiology*, 1936, **117**, 609–618.

Ramón-Moliner, E., & Nauta, W. J. H. The isodendritic core of the brain stem. *Journal of Comparative Neurology*, 1966, **126**, 311–336.

Roydes, R. L. Frontal lesions impair habituation of the head shake response in rats. *Physiology and Behavior*, 1970, **5**, 1133–1139.

Sauerland, E. K., Nakamura, Y., & Clemente, C. D. The role of the lower brain stem in cortically induced inhibition of somatic reflex in the cat. *Brain Research*, 1967, **6**, 164–179.

Scheibel, M. E., & Scheibel, A. B. Structure patterns for integrative patterns in the brain stem reticular core. In H. H. Jasper et al. (Eds.), *Reticular formation of the brain.* Boston: Little, Brown, 1958.

Scheibel, M. E., & Scheibel, A. B. The response of reticular units to repetitive stimuli. *Archives of Italian Biology*, 1965, **103**, 279–299.

Scheibel, M. E., & Scheibel, A. B. Anatomical basis of attention mechanisms in vertebrate brains. In G. Quarton, J. Melnechuk, & F. Schmitt (Eds.), *The neurosciences.* New York: Rockefeller University Press, 1967.

Shlag, J., & Challet, F. Thalamic mechanisms involved in cortical desynchronization and recruiting responses. *Electroencephalography and Clinical Neurophysiology*, 1963, **15**, 39–62.

Sharpless, S. K., & Jasper, H. Habituation of the arousal reaction. *Brain*, 1956, **79**, 655–680.

Sherrington, C. S. *The integrative action of the nervous system.* New Haven: Yale University Press, 1906.

Sokolov, Y. N. *Perception and the conditioned reflex.* (Trans. by S. W. Seydenfeld) Oxford: Pergamon Press, 1963.

Spencer, W. A., Thompson, R. F., & Neilson, D. R., Jr. Response decrement of flexion in acute spinal cat and transient restoration by strong stimuli. *Journal of Neurophysiology*, 1966, **29**, 221–239. (a)

Spencer, W. A., Thompson, R. G., & Neilson, D. R., Jr. Alternations in responsiveness of ascending and reflex pathways activated by iterated cutaneous afferent volleys. *Journal of Neurophysiology*, 1966, **29**, 240–252. (b)

SPENCER, W. A., THOMPSON, R. F., & NEILSON, D. R., JR. Decrement of ventral root electrotonus and intracellularly recorded post-synaptic potential produced by iterated cutaneous afferent volleys. *Journal of Neurophysiology*, 1966, **29**, 253–274. (c)

STEIN, L. Habituation and stimulus novelty: A model on classical conditioning. *Psychological Review*, 1966, **73**, 352–356.

STERMAN, M., & CLEMENTE, C. D. Basal forebrain inhibition. In S. Kety, E. Evarts, & H. Williams (Eds.), *Sleep and altered states of consciousness*. Baltimore: Williams & Wilkens, 1967.

THOMPSON, R. F., & SPENCER, W. A. Habituation: A model phenomena for the study of neurona substrates of behavior. *Psychological Review*, 1966, **173**, 16–43.

THOMPSON, R. F., & WELKER, W. I. Role of auditory cortex in reflex head orientation by cats to auditory stimuli. *Journal of Comparative and Physiological Psychology*, 1963, **56**, 996–1002.

VALVERDE, F. Reticular formation of the pons and medulla oblongata. A golgi study. *Journal of Comparative Neurology*, 1961, **116**, 71–100.

VALVERDE, F. Reticular formation of the albino rat's brain stem. Cytoarchitecture and corticofugal connections. *Journal of Comparative Neurology*, 1962, **119**, 25–54.

WALDRON, H., & GWYN D. Descending nerve traits in the spinal cord of the rat: I. Fibers from the midbrain. *Journal of Comparative Neurology*, 1969, **137**, 143–155.

WEBSTER, W. R. Electrophysiological and behavioral studies of habituation. Unpublished doctoral dissertation, University of Sydney, 1969.

WEINBERGER, N. M., GOODMAN, D. A., & KITZES, L. M. Is behavioral habituation a function of peripheral auditory system blockade? *Communications in Behavioral Biology*, 1969, **3**(Pt. A), 111–116.

WEINBERGER, N. M., VELASCO, M., & LINDSLEY, D. B. Effects of lesions upon thalamically induced electrocortical desynchronization and recruiting. *Electroencephalography and Clinical Neurophysiology*, 1965, **18**, 369–377.

WICKELGREN, B. G. Habituation of spinal interneurons. *Journal of Neurophysiology*, 1967, **30**, 1424–1438.

(Received October 11, 1971)

Psychological Review
1972, Vol. 79, No. 3, 245–260

CONJOINT-MEASUREMENT FRAMEWORK FOR THE STUDY OF PROBABILISTIC INFORMATION PROCESSING[1]

THOMAS S. WALLSTEN[2]

University of North Carolina at Chapel Hill

Certain assumptions are invoked, implicitly or explicitly, whenever a descriptive model of human processing of probabilistic information is built around Bayes' rule. This paper shows that the two primary assumptions are equivalent to a very general additive conjoint-measurement model. These plus an additional assumption concerning the nature of sequential effects are then proved to be equivalent to a distributive conjoint-measurement model, given a certain task restriction. This allows the three assumptions to be tested with ordinal data. It also provides a framework for the investigation of specific theoretical problems concerning the nature and determinants of composition rules and scale values. The notions of misperception, misaggregation, and response bias currently discussed in the literature may be viewed as one subset of these problems. This approach generates experiments of a sort not previously done. One of them is presented herein. The basic model was not rejected for most of the 12 subjects, and certain diagnostic properties demonstrated how it failed for the others. In addition, specific relationships concerning the scale values emerged. The paper concludes with a discussion of the theoretical prospects resulting from this development.

A question of some interest to students of decision behavior has been that of how humans alter their opinions in the light of new evidence. More specifically, investigation has been directed to the problem of the revision of subjective probability distributions over a set of mutually exclusive and exhaustive hypotheses when new information is obtained and that information is probabilistic in nature.

Mathematically, the appropriate method for such revision is given by Bayes' rule (see e.g., Edwards, Lindman, & Savage, 1963), and in much of the psychological research, this has been used as a base line against which to compare human revision of opinion (Edwards, 1968; Rapoport & Wallsten, 1972; Slovic & Lichtenstein, 1972). It is contended here that when considered as a statement about psychology, rather than mathematics, Bayes' rule carries with it strong, and as yet untested, assumptions. Certain of these, in turn, imply a very general model, of which Bayes' rule is a special case.

The theory of conjoint measurement (Krantz, Luce, Suppes, & Tversky, 1971; Krantz & Tversky, 1971) indicates how the model is to be tested and how it is to be used to obtain subjective scale values, and offers some guidelines as to isolating the locus of violations. The latter is particularly important if one's aim is to modify the basic theory so as to render it descriptive.

Previous Descriptive Use of Bayes' Rule

To introduce Bayes' rule for purposes of this paper, consider a multistage decision task in a stationary environment. One of two states, H_i, $i = 1, 2$, is in effect, with prior probabilities $P_0(H_i)$, where $P_0(H_1) + P_0(H_2) = 1$. Each H_i specifies a discrete probability distribution over m events, E_k, $k = 1, \cdots, m$, with $\sum_{k=1}^{m} P(E_k | H_i) = 1$. An E_k is sampled with replacement and observed on each trial. Let D_n stand for the data, a sequence of n observed E_k's. Bayes' rule provides the mathematical procedure for combining the prior probabilities with the sampling probabilities after n observations to obtain the posterior probabilities, $P_n(H_i | D_n)$, concerning which H_i is generating the events. The form of the rule most useful here is

---

[1] This research was supported by Grant MH-10006 from the National Institute of Mental Health to the L. L. Thurstone Psychometric Laboratory; by the National Aeronautics and Space Administration (NASA) Grant NGR 23-005-171 to the Engineering Psychology Laboratory, University of Michigan, monitored by the Ames Research Center, NASA; and by the National Institute of Mental Health Postdoctoral Fellowship 1-F02-MH-294301-01. The author thanks David H. Krantz, Amnon Rapoport, and Forrest W. Young for many extremely helpful comments and discussions, and Francis W. Irwin, Edward S. Johnson, David H. Krantz, Jay R. Levinsohn, and Amnon Rapoport for critically reading a draft of this paper.

[2] Requests for reprints should be sent to Thomas S. Wallsten, L. L. Thurstone Psychometric Laboratory, University of North Carolina, Chapel Hill, North Carolina 27514.

$$\Omega_n = L\Omega_0 \qquad [1]$$

where Ω_0, the prior odds, is the ratio $P_0(H_1)/P_0(H_2)$; Ω_n, the posterior odds, is the ratio $P_n(H_1 | D_n)/P_n(H_2 | D_n)$; and L, the likelihood ratio, is the ratio $P(D_n | H_1)/P(D_n | H_2)$. Independence of trials implies that when an additional E_k is observed, Ω_n is multiplied by L to obtain Ω_{n+1}. In other words, L for a collection of events equals the product of the values of L for the individual E_k.

Experimental Paradigm and Results

A paradigmatic experimental task (e.g., Phillips & Edwards, 1966) can be conceptualized as one in which the subjects are shown two bookbags, H_1 and H_2, each containing red and blue chips. The fractions of red and blue in Bag H_i are p_i and $q_i = 1 - p_i$, respectively. With probability $P_0(H_i)$, Bag H_i is selected and chips are randomly sampled with replacement. The subject is shown the chips, but not the bag from which they have come. His task is to give posterior probability estimates or, more usually, posterior odds estimates as to which bag is being sampled. These estimates are compared with the Bayesian posterior values, often by means of the accuracy ratio, first used by Peterson, Schneider, and Miller (1965), and defined as the ratio of $\log_{10}$ of the subjects' odds estimates to $\log_{10}$ of the Bayesian posterior odds.

Subjects have been compared with Bayes' rule over a wide range of situations, not restricted to choices between multinomial distributions. The literature has been reviewed by DuCharme (1969), Slovic and Lichtenstein (1972), and Rapoport and Wallsten (1972). The results of most of these studies are quite consistent, in that subjects' estimates are monotonic with, but conservative relative to, the Bayesian values; they are always distributed more equally among the H_i than they should be. This conservatism is greater (accuracy ratio is smaller) at more extreme posterior probabilities, for larger values of L associated with a given E_k, for a larger number of E_k per observation, and for greater numbers of observations.

Theoretical Interpretations of Conservatism

Implications of these results have been discussed by Edwards (1968), DuCharme (1969), Slovic and Lichtenstein (1972), and Rapoport and Wallsten (1972). These authors point to three interpretations of the conservatism phenomenon which have received attention in the literature. One has been termed misperception (Beach, 1966; Vlek & Van der Heijden, 1967; Wheeler & Beach, 1968), and states that subjects combine information appropriately, but misperceive the diagnosticity (i.e., the value of L) of that information. Another has been called misaggregation, and states that subjects accurately perceive the diagnosticity of information, but incorrectly aggregate it with other information. This is favored by Edwards (1968), who has developed a man-computer interaction system to overcome this presumed deficiency in actual decision-making situations (Edwards, Phillips, Hays, & Goodman, 1968). The third explanation is that subjects might perceive and aggregate information accurately, but that they have response biases against the use of extreme numbers and that this results in apparent conservatism. This was proposed by DuCharme (1969), who offered an experiment (DuCharme, 1970) in support of his position. More general notions of response artifact were previously put forward by Peterson (see Edwards, 1968) and by Vlek and Van der Heijden (1967).

Each of these positions is incomplete as presently constituted, as the various authors themselves indicate. For example, DuCharme (1969) and Green (1968) point out that neither the misperception nor the misaggregation hypotheses offer any a priori reason why the processing errors should always be in the direction of conservatism. Furthermore, every time an experiment is done to distinguish between the two possibilities, it becomes less obvious where in the system perception stops and aggregation begins (Peterson, DuCharme, & Edwards, 1968; Wheeler & Beach, 1968; and the comments on these by Edwards, 1968). This is mainly because a theory relating perception and aggregation of data, respectively, to subjective odds has not been sufficiently developed.

In addition, investigation of the former two hypotheses has depended critically on the assumption that subjects' estimates are identity measures of their subjective posterior odds or probabilities (see Rapoport & Wallsten, 1972). The response bias explanation depends on the slightly weaker assumption that the subjective posterior values are more extreme than the estimated posterior values.

The validity of these assumptions is difficult to assess in the absence of a theoretical framework, as has been pointed out previously in a somewhat related situation (Wallsten, 1971), and discussed in Rapoport and Wallsten (1972). From the present point of view, the problem is twofold. First, a theory about how probabilistic information is processed is required to relate the subjective posterior odds to observable behavior. Second, presuming the theory is not empirically rejected, it can be used with data to scale the subjective dimensions or to evaluate particular assumptions about the nature of those scale values.

It is immediately apparent that the initial facet of the problem, that of constructing a theory, is of considerable importance, and that when it is satisfactorily resolved, the latter pales into relative insignificance. This is simply because whenever data and an appropriate theory are available, it is unnecessary to make a priori assumptions about scale values, and whenever data and theory are not available, it is impossible to assess any assumptions.

Of course one model for probabilistic information processing is Bayes' rule. In fact, it has been suggested that at least a first approximation to a descriptive model may be had by raising L in Equation 1 to a power c, $0 < c \leq 1$ (Edwards, 1968; Peterson et al., 1968; Schum & Martin, 1968). Questions of misperception and misaggregation, then, would be questions concerning why c is less than one, while questions of response bias would assume that c equals one.

Basic Assumptions

A descriptive model built around Bayes' rule carries with it two important independent assumptions which, perforce, are incorporated in all the theoretical and empirical work discussed above. One, the likelihood ratio principle, states that the subjective diagnosticity of an event (or set of events) depends on the ratio of the subjective likelihood of the event(s) given one H_i to that given the other. The other assumption, the multiplying principle, states that the subjective diagnosticity of a set of events equals the product of the subjective diagnosticities of these events taken singly.

The validity of each of these assumptions is unknown despite their crucial positions. This has been stated perhaps most forcefully by Green (1968), who pointed out that a linear function (in this case the log of Equation 1 with L^c substituted for L) can often account for most of the variance generated by a more complex function and, in the absence of prior theoretical notions, can easily be falsely accepted. Thus, although it may provide a reasonable description of the subjects' responses, it is an open question whether it provides a valid explanation of the processes which led to those responses.

The assumptions which underlie the previous work may now be summarized: (*a*) the likelihood ratio principle, (*b*) the multiplying principle, and (*c*) the assumption that subjects' probability or odds estimates are identity measures of their subjective probabilities. Occasionally *c* is replaced by (*d*) the assumption that subjects' responses are conservative relative to their subjective probabilities. In addition, various assumptions have been made about what constitutes perception rather than aggregation of the information, and vice versa.

Assumptions *a* and *b* are the cornerstones. They represent the main statements concerning how humans revise their opinions in the light of uncertain evidence, or how they process probabilistic information. These principles are also embedded in other, related psychological work, such as optional stopping (Edwards, 1965; Rapoport & Burkheimer, 1971) and signal detection theory (Green & Swets, 1966). Their validity has been investigated to some degree in the latter case (Green & Swets, 1966, pp. 101–107; Swets & Birdsall, 1967; Ulehla, Canges, & Wackwitz, 1967; Ulehla, Halpern, & Cerf, 1968).

GENERAL MODEL AND CONJOINT-MEASUREMENT ANALYSIS

From a psychological point of view, Assumption *a* in its weakest form is no more than a composition principle (Krantz & Tversky, 1971). That is, a rule specifying the manner in which the dimensions of a stimulus are combined by the subject. To make this point clear consider a set of data, D_n, presented to the subject, following which he revises his opinion concerning two mutually exclusive and collectively exhaustive states. To facilitate the development, the states will be denoted X and Y, respectively, instead of H_i. The likelihood ratio principle states that D_n has two relevant dimensions, $P(\mathrm{D}_n|\mathrm{X})$ and $P(\mathrm{D}_n|\mathrm{Y})$. The subjective posterior odds are a monotonic function of the ratio of the subjective correlates of these dimensions. The subject emits a response with at least ordinal properties, which itself is monotonically related to the posterior odds. Expressed as an equation and indexing the X and Y, respectively, Assumption *a* may be interpreted as

$$R(\mathrm{X}_i, \mathrm{Y}_j|\mathrm{D}_n) = f\left[\frac{\Phi_1(\mathrm{D}_n|\mathrm{X}_i)}{\Phi_2(\mathrm{D}_n|\mathrm{Y}_j)}\right], \qquad [2a]$$

where the left-hand member of the expression is the subject's response, favoring X_i rather than Y_j given D_n. The Φ_m, $m = 1, 2$, are real valued functions representing the subjective correlates of the conditional probabilities, and f is increasing in Φ_1 and decreasing in Φ_2. To be completely general, it is necessary to specify two functions, Φ_1 and Φ_2, since there is no a priori reason to believe that subjects treat the two relevant dimensions of D_n equivalently. In all that follows, X will refer to that state about whose likelihood the subject is explicitly responding, and Y will refer to the other.

Thus, the likelihood ratio principle is equivalent to a multiplicative conjoint-measurement model, with the numerator and denominator of the ratio as factors within the structure. If the multiplicative model is shown to be invalid, the principle must be rejected at least for that situation.

It is important to be clear concerning the two factors in the model. Manipulating these two variables does not require manipulation of the physical structure of D_n. Rather, for a *given* D_n, the values of $P(\mathrm{D}_n|\mathrm{X}_i)$ and $P(\mathrm{D}_n|\mathrm{Y}_j)$ are varied by varying the X_i and Y_j. It may be the case that the Φ_m are influenced by the physical nature of the sample, but that is a separate issue to be investigated given that the validity of Equation 2a is not rejected.

Notice that, disallowing negative values of the Φ_m, Equation 2a is formally identical to an additive model. That is,

$$R(\mathrm{X}_i, \mathrm{Y}_j|\mathrm{D}_n) = g[\phi_1(\mathrm{D}_n|\mathrm{X}_i) - \phi_2(\mathrm{D}_n|\mathrm{Y}_j)], \qquad [2b]$$

where $\phi_m = \log\Phi_m$, $m = 1, 2$. In view of the close correspondence in meaning between the Φ_m and the $P(\mathrm{D}_n|\mathrm{H}_i)$, the assumption of nonnegative Φ_m seems reasonable. The distinction between the additive and multiplicative versions will be returned to below.

Turning now to Assumption *b*, it too, in its weakest form, is a composition rule. This has been pursued by Shanteau (1970), within the framework of functional measurement, and by Lathrop.[3] Consider the sequential presentation (in time or space) of the elements, E_k, of a set of discrete data, following which the subject revises his opinion concerning X_i and Y_j. According to the multiplying principle, the subjective diagnostic values of the E_k multiply to form the diagnostic value of the sample. The subjective posterior odds and, therefore, the response are monotonically related to this product. Expressed as an equation,

$$R(X_i, Y_j | D_n) = f[\Psi_1(E_{k1})\Psi_2(E_{k2}) \cdots \Psi_n(E_{kn})], \quad [3a]$$

where $\Psi_h(E_{kh})$, $h = 1, \cdots, n$, are real valued functions representing the subjective diagnostic value of event E_k on the hth trial, and f is increasing in Ψ_h. Thus, trivially, the multiplying principle is equivalent to a multiplicative conjoint-measurement model, with positions in the data sequence as factors and the E_k as levels within the factors.

Shanteau (1970) proposed a weighted averaging model of inference as an alternative to the Bayesian. This, of course, is a special case of a general additive model. However, Equation 3a could as well have been written as an additive model. That is,

$$R(X_i, Y_j | D_n) = g[\psi_1(E_{k1}) + \cdots + \psi_n(E_{kn})], \quad [3b]$$

where $\psi_h = \log \Psi_h$, $h = 1, \cdots, n$, assuming Ψ_h is always nonnegative. This assumption seems reasonable given the close correspondence in meaning between the subjective diagnosticity, Ψ_h, and the likelihood ratio, L in Equation 1.

However, additive theories would appear to be quite different psychologically from multiplicative theories, and it is an important question whether they can be distinguished empirically. It is argued that the answer is no in the case of Equations 2a and 2b and 3a and 3b, respectively, but the argument cannot be made without introducing the notion of sign dependence, which follows from a multiplicative composition rule. It is formally developed by Krantz and Tversky (1971) and Krantz et al. (1971), who demonstrate that this is the only empirical distinction between the two models.

Sign dependence holds if two or more factors multiply and at least one of them can be divided into three subsets containing elements with positive, zero, and negative values, respectively. Then the rank orders induced on the responses by the levels of the other factor(s) given a positive element of the first will be reversed given a negative element, and made degenerate given a zero element.

Evidence of sign dependence in the data is diagnostic of a multiplicative rather than an additive model. The converse does not hold, since the experiment simply may not have employed a sufficiently wide range of levels of a factor to have included subjective values of opposite signs or of zero. Thus, whenever an additive model is supported by the data, so is the corresponding multiplicative model.

It is now further argued that in the present context, sign dependence can never be diagnosed conclusively; therefore, it will remain impossible to reject Equations 2b or 3b in favor of Equations 2a or 3a, respectively. It is maintained by Krantz and Tversky (1971) and Krantz et al. (1971) that sign dependence cannot be diagnosed unless at least *any* two of the positive, zero, or negative subsets are nonempty. Thus, if it is correct that Ψ_h and Φ_m must be nonnegative for all h and m, then to observe sign dependence, it is necessary that for at least one h, $\Psi_h = 0$, and for at least one m, $\Phi_m = 0$. The effect of these values with a multiplicative process would be to induce a degenerate ordering over the dependent variable. However, very large scale values will have the same effect with an additive model. It follows, therefore, that although a degenerate ordering may lead one to suspect a multiplicative process (if the requisite scale values under an additive model seem unlikely), it cannot be considered criterial. Thus, the multiplicative can be distinguished from the additive model only when at least the positive and negative subsets are nonempty, which in the present case will never occur.

---

[3] R. G. Lathrop. Riskless human decisions: A tutorial and a model. Unpublished manuscript, 1970.

The immediate theoretical effect of this is that in the absence of additional assumptions, the question whether subjects behave as if they multiply or add the subjective diagnosticities of events is unanswerable. Similarly, the question whether these diagnostic values depend on ratios or differences of subjective conditional likelihoods is also unanswerable. Of course, different theoretical formulations may be more or less useful, depending on other considerations, but the "crucial experiment" will remain impossible.

The multiplying principle, or Equation 3b, is concerned with the relation between the subjective diagnosticities of events and of samples, but it says nothing about what gives rise to the diagnostic values in the first place. Conversely, the likelihood ratio principle, or Equation 2b, is a statement concerning the origins of the diagnostic values, but it does not comment on the relations between the diagnostic values of events and of samples. Each could be tested independently using an experimental design dictated by considerations of conjoint measurement. The two principles may be combined into a single model by considering each of Equations 2b and 3b for a sample consisting of one event E_k. From this it follows that

$$\Psi_1(E_k) = \phi_1(E_k | X_i) - \phi_2(E_k | Y_j). \quad [4]$$

Strictly speaking, the terms on the right-hand side should be subscripted to reflect the position number. Developments in the Appendix to this paper, however, obviate this. The logic which led to Equation 4 may be pursued when a second event is added to the sample, etc., to obtain a very general (and unwieldy) additive model.

Instead, it is useful to consider the special case in which the sequence of data consists of n replicates of the same event E. Using k now to index the number of replicates, n_k, it is shown in the Appendix that Equations 2b and 3b (or 2a and 3a) plus an additional assumption (e) are valid if and only if

$$R(X_i, Y_j | n_k E) = g\{\phi_3(n_k) \times [\phi_1(E | X_i) - \phi_2(E | Y_j)]\}, \quad [5]$$

where $\phi_3(n_k)$ is a real valued function representing the subjective value of n_k replicates of the event E. The function g is increasing in ϕ_3 and ϕ_1, and decreasing in ϕ_2. Assumption e states that the effect of position is independent of the conditional likelihoods which occur at *that* position.

Thus, in the special case being considered, the likelihood ratio principle together with the multiplying principle and Assumption e are equivalent to a distributive conjoint-measurement model, with the following three independent variables as factors within the structure: $P(E|X_i)$, $P(E|Y_j)$, and n_k. If the distributive model is empirically valid, ϕ_1 and ϕ_2 can be scaled uniquely up to linear transformations of $\phi_1' = a\phi_1 + b$ and $\phi_2' = a\phi_2 - b$, $a > 0$, and ϕ_3 can be scaled uniquely up to a ratio transformation (Krantz et al., 1971). If the model is empirically rejected, certain diagnostic properties may indicate which of the assumptions may be in error.

Equation 5 suggests a factorial experiment in which n_k, X_i, and Y_j are varied. Recall that manipulating X_i and Y_j results in manipulation of the numerator, $P(E|X_i)$, and the denominator, $P(E|Y_j)$, respectively, of the likelihood ratio for the event E. The axioms of conjoint measurement are formulated in this framework; those which are necessary and constitute the diagnostic properties can be tested using only ordinal aspects of the data (Krantz & Tversky, 1971; Krantz et al., 1971). Another evaluation more in the spirit of functional measurement (Anderson, 1970) and based on a necessary and sufficient condition is also possible.

AN EXPERIMENT

In reporting the experiment, it is convenient to let the set of $P(E|X_i)$ be collectively denoted by **N** (for numerator), the set of $P(E|Y_j)$ by **D** (for denominator), and the set of n_k by **n**. The study involved

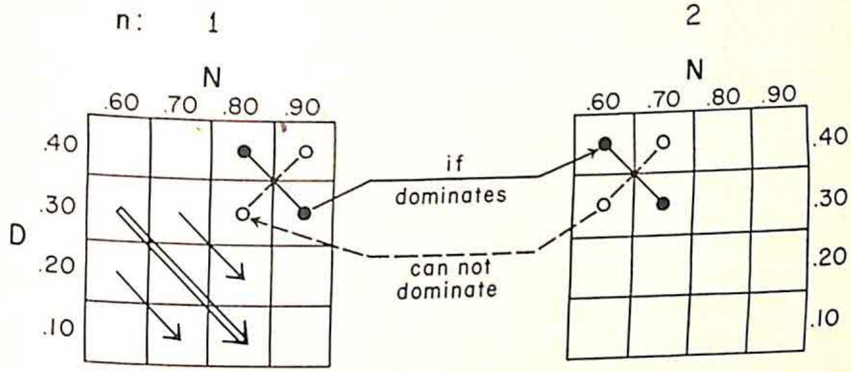


FIG. 1. Illustration of experimental design showing four levels of **N**, four of **D**, and two of **n**. The arrows, etc., illustrate cancellation axioms.

a 4 × 4 × 2 factorial design to fill in the **D** × **N** × **n** matrix of Figure 1.

Method

Procedure and apparatus. On every trial, the subject observed events to decide which of two a priori equally likely binomial distributions was in effect. The distributions were physically represented by spinners, labeled 1 and 2, respectively. They were 20 centimeters (cm.) in diameter, mounted in black Plexiglas bases, and each was divided into a green and a silver radial sector. The events, therefore, were the occurrences of the spinner stopping over green or silver. The relative areas of the sectors were adjusted before each trial to change the distributions. The bases were calibrated on the back with no markings visible to the subject, and during most of the experiment the bases stood upright, side by side, so as to be easily viewed.

The subject was truthfully told that prior probabilities were equal and that a computer had randomly selected one of the two spinners currently displayed and had simulated its being spun. Thus, an experimental trial proceeded as follows: The two spinners were set in full view of the subject, and he was given a slip of paper on which the results of the spins (the events) were listed. He wrote on the slip the number of the spinner he thought more likely to have been used and gave a probability estimate of being right. The subject returned the slip and was informed of the correct alternative. The spinners were reset and another trial began.

Estimates were collected in two ways. The estimate group marked a 20-cm.-long scale calibrated in units of .01 from .50 to 1.00. The line group drew a line in a 20-cm.-long closed box, which was later measured by superimposing the calibrated scale. All analyses were performed on the responses, r, transformed to log odds, $\log r/(1-r)$.

Design. The 10 decision problems used are indicated by the x's in Table 1. For example, the x in row 2, column 3, indicates that the subject had to decide whether the spinner with 30% green and 70% silver generated the observations or whether one with 80% green and 20% silver did. Spinner 1 in the experiment corresponded to Spinner A in the table and Spinner 2 corresponded to Spinner B for a random half of the trials. This was reversed for the other half. Half of the time the subject was required to decide on the basis of observing one event in order to generate data for the left-hand plane in Figure 1, that is, for $n_1 = 1$. The other half of the time he was given two observations to generate data for the right-hand plane, that is, for $n_2 = 2$.

Now it is shown that the problems in Table 1 satisfy the demands of the **D** × **N** planes of Figure 1. Consider again the previous example, row 2, column 3, of Table 1, with a single observation. If that event were green, thereby favoring the .8–.2 spinner, and if the subject's response also favored that spinner, the response would be entered in Cell (2,3) of the left-hand plane. That is, the cell with the row variable $P(E|Y_2) = .3$, and the column variable $P(E|X_3) = .8$. If the event were silver and the subject's response favored the .3–.7 spinner, the response would be entered in Cell (3,2) of the plane, that is, the cell with the row variable $P(E|Y_3) = .2$, and the column variable $P(E|X_2) = .7$. In this manner, the 10 problems yield data to fill the left-hand plane of Figure 1.

In the same way, the right-hand plane was filled when decisions were based on two observations in-

TABLE 1

EXPERIMENTAL DESIGN

Spinner B	Spinner A			
	.6–.4	.7–.3	.8–.2	.9–.1
.4–.6	x	x	x	x
.3–.7		x	x	x
.2–.8			x	x
.1–.9				x

Note.—The first number in each pair is the proportion of green on the spinner; the second number is that of silver. The settings which were paired together are indicated by an x.

TABLE 2

GROUP MEAN LOG ODDS ESTIMATES CONVERTED TO ODDS, AND ACCURACY RATIOS FOR THE SYMMETRICAL PROBLEMS

Spinner	Spinner							
	$n_1 = 1$				$n_2 = 2$			
	.6–.4	.7–.3	.8–.2	.9–.1	.6–.4	.7–.3	.8–.2	.9–.1
	Estimate group							
.4–.6	1.67	1.80	2.50	3.06	2.10	2.83	3.68	4.68
.3–.7	2.22	2.51	2.74	3.27	2.73	4.33	4.79	6.71
.2–.8	3.13	3.38	3.46	4.38	5.05	4.70	6.62	8.40
.1–.9	4.80	4.07	5.64	7.15	10.00	8.93	9.68	14.69
Accuracy ratio of diagonal entry	1.26	1.08	.896	.895	.913	.865	.682	.611
	Line group							
.4–.6	1.82	2.40	2.60	3.33	2.48	3.54	5.35	6.08
.3–.7	2.31	3.06	3.03	3.02	3.91	4.90	5.34	6.39
.2–.8	4.11	3.88	3.55	4.12	5.17	5.71	6.09	11.24
.1–.9	5.35	6.18	6.02	7.44	12.72	17.08	11.50	20.09
Accuracy ratio of diagonal entry	1.48	1.32	.914	.913	1.12	.938	.652	.683

Note.—Entries other than accuracy ratios are the values x representing the odds x:1 in favor of the column spinner when paired with that along the row.

stead of one. However, whenever the two events included one green and one silver, the subject was given this information, just as on all trials, but these responses were not analyzed. Note that the particular cell in which a response is entered depends on a random event; therefore, the number of entries per cell will not necessarily be equal.

Subjects' responses were discarded when they rated the mathematically more likely alternative as less likely. This was necessary to maintain the correspondence between the ϕ_m, $m = 1, 2$, and the $P(\mathrm{E} \mid \mathrm{X}_i)$ and $P(\mathrm{E} \mid \mathrm{Y}_j)$, respectively. Only 1.2% of the responses destined for the $\mathbf{D} \times \mathbf{N} \times \mathbf{n}$ matrix fell in this category, and of these, all the responses but two were from subjects for whom subsequent analyses tended to reject the model.

Subjects. There were six male subjects per group, all University of Michigan undergraduates paid $1.75 per hour, each run individually for two sessions of about 90 minutes apiece. The first session included a practice block of 20 trials, in which each problem appeared twice in random order, once with one event and once with two. This was followed by four similar blocks for data. Session 2 consisted of six experimental blocks, making a total of 200 experimental trials, 20 of each problem, half with one event and half with two.

Results and Discussion

Table 2 shows the group means of the transformed estimates, converted back to odds. The main analyses were performed on individual subject's responses, but Table 2 is a useful summary of the raw data and serves to provide a comparison with other studies.

Most experiments have run subjects on symmetrical problems only, that is, those in the diagonals of the table; the present results may be compared to others in terms of accuracy ratios for those problems. These are given in Table 2. It is evident that on the average the subjects were increasingly conservative as n_k increased from 1 to 2, and as L increased, thereby duplicating the essential findings of others. Additionally, subjects in the line group were consistently more extreme than those in the estimate group, as also reported by Lathrop, Jones, and Keiser (1969). Thus, in terms of these gross measures, the present data are compatible with the literature.

Turning now to the composition rule analyses for individual subjects, it is convenient to discuss a necessary and sufficient condition for the distributive model before discussing those which are only necessary. To facilitate this let

$$\theta(\mathrm{X}_i, \mathrm{Y}_j, n_k) = \phi_3(n_k)[\phi_1(\mathrm{E}|\mathrm{X}_i) - \phi_2(\mathrm{E}|\mathrm{Y}_j)]. \quad [6]$$

Then the distributive model (Equation 5) is satisfied if and only if there exist ϕ_m, $m = 1, 2, 3$, such that for all i, j, k,

$$R(\mathrm{X}_i, \mathrm{Y}_j|n_k\mathrm{E}) > R(\mathrm{X}_{i'}, \mathrm{Y}_{j'}|n_{k'}\mathrm{E}) \text{ iff } \theta(\mathrm{X}_i, \mathrm{Y}_j, n_k) > \theta(\mathrm{X}_{i'}, \mathrm{Y}_{j'}, n_{k'})$$

and $\quad [7]$

$$R(\mathrm{X}_i, \mathrm{Y}_j|n_k\mathrm{E}) = R(\mathrm{X}_{i'}, \mathrm{Y}_{j'}|n_{k'}\mathrm{E}) \text{ iff } \theta(\mathrm{X}_i, \mathrm{Y}_j, n_k) = \theta(\mathrm{X}_{i'}, \mathrm{Y}_{j'}, n_{k'}).$$

Although Condition 7 is easily expressed, its evaluation is not trivial. One method, developed for any of the simple polynomial models[4] is based in the present case on searching for ϕ_m, $m = 1, 2, 3$, to minimize

$$\sum_{ijk} [\hat{R}(\mathrm{X}_i, \mathrm{Y}_j, n_k) - \theta(\mathrm{X}_i, \mathrm{Y}_j, n_k)]^2, \quad [8]$$

where

$$\hat{R}(\mathrm{X}_i, \mathrm{Y}_j, n_k) \geqq \hat{R}(\mathrm{X}_{i'}, \mathrm{Y}_{j'}, n_{k'}) \text{ iff } R(\mathrm{X}_i, \mathrm{Y}_j|n_k\mathrm{E}) \geqq R(\mathrm{X}_{i'}, \mathrm{Y}_{j'}|n_{k'}\mathrm{E}),$$

for all i, j, k. In other words, $\hat{R}$ is a weak monotonic transformation of R, and the effect of minimizing Expression 8 is to use a least-squares procedure on ordinal data in a manner similar to that developed by Kruskal (1965).

Monotonicity constraints were imposed on the ϕ_m. That is, ϕ_1 was weakly monotonically related to $P(\mathrm{E}|\mathrm{X}_i)$, etc. But this is innocuous; none of the ϕ_m were forced to any shape nor to any relation with other ϕ_m.

If the minimum for Expression 8 is found to be zero, then the terms being summed are each zero, since they are squares and, therefore, nonnegative. In that case, $\hat{R} = \theta$, for all i, j, k, and θ, perforce, is (at least weakly) monotonically related to R. It is a statistical question, however, whether Condition 7 is satisfied, since the solution may be a degenerate one in which

$$\theta(\mathrm{X}_i, \mathrm{Y}_j, n_k) = \theta(\mathrm{X}_{j'}, \mathrm{Y}_{j'}, n_{k'}),$$

for many i, j, k.

Alternatively, the minimum may be greater than zero, in which case monotonicity between θ and R is certainly not perfect. An appropriate (but not very powerful) statistic in either case is Kendall's tau (τ) to evaluate the rank-order correlation between θ and R.

Strictly speaking, the present analysis did not minimize Expression 8, since the computer was programmed to stop searching the parameter space when a complete pass through all the $\hat{R}$ or all the θ decreased the sum by less than .01. The effects of this are unknown, but probably negligible.

The results of the minimization analyses are displayed, separately for each subject, in Table 3 and Figure 2. The scales, ϕ_1 and ϕ_2, are graphed as a function of $\log P(\mathrm{E}|\mathrm{X}_i)$ and $\log P(\mathrm{E}|\mathrm{Y}_j)$, respectively. These are interval scales sharing two degrees of freedom between them and allowing two values,

$$\phi_1(\mathrm{E}|\mathrm{X}_1) = \log P(\mathrm{E}|\mathrm{X}_1) = -.5108$$

and

$$\phi_2(\mathrm{E}|\mathrm{Y}_1) = \log P(\mathrm{E}|\mathrm{Y}_1) = -.9163,$$

to be set. These values are circled in Figure 2. The values for $\phi_3(n_2)$ are shown in column 4 of the table. Since ϕ_3 is a

TABLE 3

RESULTS OF MINIMIZATION ANALYSIS BASED ON CELL MEANS

Subject	τ	% agree	$\phi_3(n_2)$	Slope **N**	Slope **D**
1E	.750	87.5	1.56	1.11	.607
2E	.630	81.5	1.16	.767	.439
3E	.734	86.7	1.58	1.36	.956
4E	.817	90.9	1.74	1.58	.911
5E	.855	92.8	1.75	1.15	.613
6E	.589	79.5	1.43	1.18	.406
1L	.601	80.0	1.55	1.46	.824
2L	.794	89.7	1.55	1.14	.998
3L	.810	90.5	1.56	1.23	.822
4L	.593	79.7	2.44	.948	.322
5L	.786	89.3	1.50	1.23	.456
6L	.786	89.3	1.45	.996	.604

---

[4] T. S. Wallsten. An ordinal method for the simultaneous evaluation and parameter search of simple polynomial models (tentative title). In preparation.

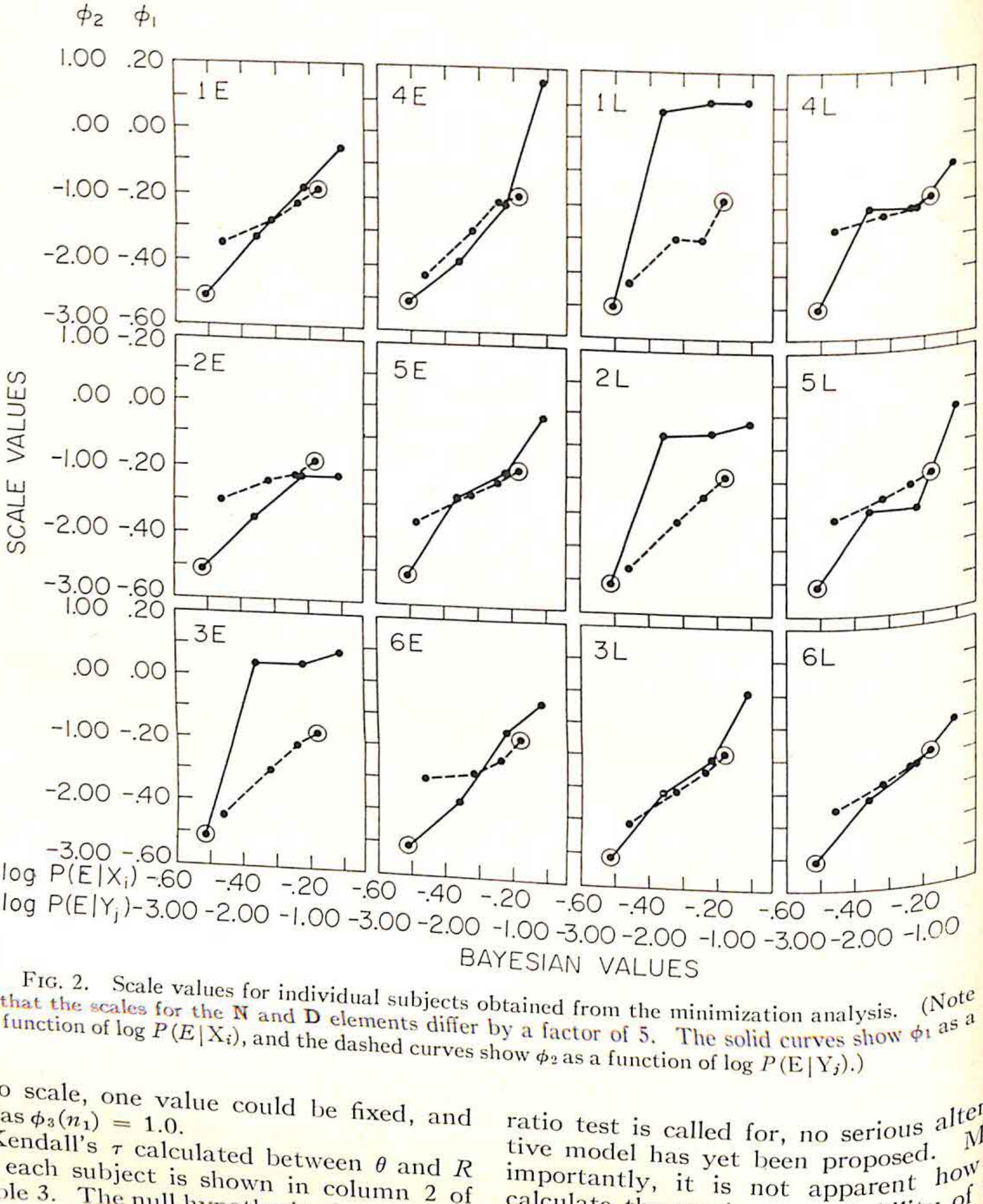


FIG. 2. Scale values for individual subjects obtained from the minimization analysis. (Note that the scales for the **N** and **D** elements differ by a factor of 5. The solid curves show ϕ_1 as a function of log $P(E|X_i)$, and the dashed curves show ϕ_2 as a function of log $P(E|Y_j)$.)

ratio scale, one value could be fixed, and it was $\phi_3(n_1) = 1.0$.

Kendall's τ calculated between θ and R for each subject is shown in column 2 of Table 3. The null hypothesis of no relation between the rank orders of θ and R can be rejected in each case ($p < .001$). But, clearly, this does not imply that the distributive model is valid, since the significant τ values may have occurred because the model correlates highly with a different (as yet unspecified) composition principle which more accurately models the processing of the stimulus dimensions. Although it appears as if a likelihood ratio test is called for, no serious alternative model has yet been proposed. More importantly, it is not apparent how to calculate the maximum probability of obtaining a certain value of τ given a particular model.

A different tack was pursued here, involving a transformation of τ. Considering θ and R each to induce an ordering on the 32 elements in the $\mathbf{D} \times \mathbf{N} \times \mathbf{n}$ matrix, one may perform pair comparisons of the elements within each ordering. There are 496 such pairs in each case, and τ is a measure of agreement between the two sets. Another measure is the simple percentage of agreement, which is shown in column 3 of Table 3. These are generally quite high and appear to cluster around 80% for four subjects, 2E, 6E, 1L and 4L (where E = estimate group and L = line group), and 90% for the others. This might be taken as evidence to reject the model for those four subjects. It may be added here that these are the subjects whose responses occasionally favored the mathematically less likely alternative. There were eight such responses for Subject 2E, three for 6E, four for 1L, and nine for 4L. (Subjects 4E and 2L also had one such response apiece.)

Before inquiring into possible causes of the model's failure in the four cases, it is instructive to look at the scale values, particularly with the remaining subjects. Note that in each case, $\phi_3(n_2)$ is less than 2. Thus, the diagnosticity of two formally identical events was considerably less than that of one. This is consistent with the finding of increased conservatism with larger samples of data, and indicates that the latter phenomenon is due to factors more basic than response bias. However, the psychological processes responsible for the present finding are not apparent, and further investigation is necessary to uncover them.

Now look at the plots of ϕ_1 and ϕ_2 in Figure 1. The general slope of ϕ_1 is greater than that of ϕ_2 in all cases. Obviously not all the plots are linear, but for descriptive purposes, the slopes of the best-fitting straight lines are shown in the last two columns of Table 3. Since the two scales are unique up to a common multiplicative constant, the ratio of the two slopes is constant under all permissible transformations.

Ideal processing of the conditional likelihoods entails identical slopes for the two scales. At least three explanations exist for the discrepancy, none of which can currently be ruled out. One is that subjects inherently distort likelihood ratios away from unity. This might be interpreted as an aversion toward uncertainty, but that would appear to be at variance with the values for $\phi_3(n_2)$.

Another interpretation is that subjects have an attention bias, causing them to weight one dimension relatively heavier than another. Thus, they may have attended more strongly to the numerator because it was associated with the spinner which was more likely correct. Factors, such as payoffs, could be imposed on the task to shift the bias.

A third explanation is that the differential slopes artifactually resulted from the particular values on the abcissa. One form this argument might assume is that subjects were taking differences instead of ratios of conditional probabilities and, therefore, the scale values should have been plotted as a function of $P(\mathrm{E}|\cdot)$ instead of $\log P(\mathrm{E}|\cdot)$. However, the effect of this would be to render the ϕ_2 slope greater than the ϕ_1 slope in all cases, thus shifting the focus of the problem, but not eliminating it.

Another form of the argument is that the conditional likelihoods are irrelevant to the subjects, and any interpretation based on them is meaningless. This, of course, is possible, but seems unlikely in the present case. The conditional likelihoods were not just abstract numbers; they were measures of a physical characteristic of the stimuli, that is, relative areas, to which most of the subjects responded appropriately at an ordinal level.

The necessary axioms are formally presented and their diagnostic potential completely discussed in Krantz and Tversky (1971). They will be informally introduced here without proof, but they all follow from appropriate applications of Equation 5. Whereas the minimization analysis was performed on cell means, the axiom analyses were performed on rank orders obtained by using the Mann-Whitney U test to compare each cell with each of the others. Thus, the error theories incorporated in the two analyses were different.

The distributive model contains an additive component, which entails the two necessary conditions of independence and double cancellation. These were evaluated separately within each of the n_k planes of Figure 1. Independence implies that the effect of the column variables, **N**, on the

TABLE 4

RESULTS OF AXIOM ANALYSES USING RANKS BASED ON MANN-WHITNEY U TESTS

Subject	Additive						Distributive	
	$n_1 = 1$			$n_2 = 2$				
	W:N;D	W:D;N	CV	W:N;D	W:D;N	CV	τ:N × D;n	CV
1E	.900	1.00	0	.814	1.00	0	.917	0 (3)
2E	.775	.925	0	.592 (*ns*)	.925	0	.667	0 (3)
3E	.785	1.00	0	.814	1.00	0	.733	0
4E	.737	.900	0	.825	.939	0	.750	0
5E	.737	1.00	0	.825	.675	0	.617	0
6E	.164 (*ns*)	.111 (*ns*)	6	.825	.675	0	.167 (*ns*)	0 (8)
1L	.447 (*ns*)	.825	0	.625	.678	0	.733	2 (3)
2L	.095 (*ns*)	.925	0	.481 (*ns*)	.850	0	.667	0
3L	.401 (*ns*)	.814	0	.700	.981	0	.750	0
4L	.276 (*ns*)	.813	9	.327 (*ns*)	.639 (*ns*)	4	.033 (*ns*)	0
5L	.650	1.00	0	.675	.712	0	.733	0
6L	.096 (*ns*)	1.00	0	.853	.968	0	.617	0

Note.—Correlations not significantly different from zero marked *ns* ($p > .05$). CV = cancellation violations.

cell entries is independent of the level of the row variables, **D**, and vice versa. Thus, for example, if **N** is independent of **D**, the rank order of the cells over the four levels of **N** is the same at $P(\mathrm{E}|\mathrm{Y}_1)$ as it is at $P(\mathrm{E}|\mathrm{Y}_2)$, etc. This was checked by using Kendall's coefficient of concordance, W, to determine the correlation of rank orders within rows and within columns, respectively. The results in the former case, testing whether **N** is independent of **D**, are shown in columns 2 and 5 of Table 4, each labeled W: **N**; **D**. The results in the latter case, testing whether **D** is independent of **N**, are shown in columns 3 and 6 of the table.

Double cancellation can be tested in as many as all of the sixteen 3 × 3 matrices of each plane, and is illustrated within the lower left corner of the left plane of Figure 1. If the rank ordering of responses is as indicated by the single arrows, then the ordering shown by the double arrows must be observed. The number of violations within each plane out of 16 tests is given in columns 4 and 7 of Table 4.

The other necessary conditions for the distributive model, which are subject to empirical verification, are joint independence, distributive cancellation, and joint sign dependence. The first two were tested across the n_k matrices of Figure 1.

Joint independence implies that the joint effect of the **N** × **D** variables on the cell entries is independent of the level of **n**, and conversely. The former was evaluated by calculating Kendall's τ between the n_1 and n_2 planes, the results of which are shown in column 8 of Table 4. The latter was tested by ranking the n_1 and n_2 cells at each of the 16 **N** × **D** levels. Agreement was perfect for seven subjects; almost perfect for four others, 1E, 2E, 1L, and 3L; and for one subject, 6E, 11 rankings were in one direction and 5 were in the other.

Distributive cancellation should hold in all pairs formed by taking a 2 × 2 minor from each of the **N** × **D** planes. It is illustrated in Figure 1 and shows that if a diagonal of one minor dominates a diagonal of the other, the reverse domination cannot hold. The number of violations is shown in the last column of Table 4. It is based on 1,296 tests except where parentheses appear, containing the number of cases in which the antecedent conditions were not met and, therefore, no test was possible.

Joint sign dependence is defined analogously to sign dependence, except that a single factor is replaced with a joint factor, **N** × **D** in the present case. It was not tested here, since the experiment did not employ the proper factor levels. It is not unreasonable to suppose, however, that had n_k, $P(E|X_i)$ or $P(E|Y_j)$ equaled zero, degenerate orderings would have been observed.

Recall, now, that the minimization analysis was interpreted to reject the model for Subjects 2E, 6E, 1L, and 4L. It is apparent from the axiom analyses that the model was totally inadequate for Subjects 6E and 4L. Failures of independence and double cancellation show that the likelihood ratio principle was violated, and failures of joint independence demonstrate that either or both of the multiplying principle or Assumption *c* was in error.

It appears that Subjects 2E and 1L, however, violated only the likelihood ratio principle. The subjective diagnosticities of the two events multiplied, or summed, regardless of how they arose in the first place.

But the axiom analyses yield the same conclusion for Subjects 2L, 3L, and 6L for whom the model originally was not rejected. Of course, the two analyses incorporate different error theories and, furthermore, the point of rejection is arbitrary, especially, it may be argued, with the minimization analysis. Nevertheless, it is problematic which conclusions to accept. It may be added that when all the analyses agree, that strongly suggests that the psychological rather than the error theory is responsible.

One way out of the dilemma in the present case is to assume that Subjects 2L, 3L, and 6L did not, in fact, violate the likelihood ratio principle. Rather, the between-cell variability of the entries in a given row was so small that inversions in rankings occurred as a matter of chance, whereas in the case of Subjects 2E and 1L, they occurred in spite of large between-cell variability. Partial support for this is obtained by looking at the standard deviations of the cell entries about their row means where independence failed. Using medians as the cell entries (since they, like the Mann-Whitney *U*, are insensitive to distance), the mean of the four standard deviations for the rows was .214 and .302 for Subjects 2E and 1L, respectively. It was .043 and .093 for the n_1 and n_2 planes of Subject 2L, .093 for Subject 6L, but .276 for Subject 3L. Similar results obtained when the calculations were on means rather than medians. Thus, it is still uncertain whether the model failed for Subject 3L.

CONCLUDING REMARKS

The experiment was illustrative on a number of points. Not the least of these is the fact that despite the absence of an explicit error theory, the conjoint-measurement technique did lead to some firm conclusions, particularly in regard to the ϕ_m. Doubt along these lines has been expressed by Zinnes (1969). Nevertheless, it is also true that had an error theory been available, even more might have been inferred from the results. Thus, referring back to Table 4, it was possible to reject the axioms in some cases, but strictly speaking they can be accepted only if the rank correlations are perfect or there are zero cancellation violations. Similarly, the necessary and sufficient Condition 7 is satisfied only when $\tau = 1.0$. Thus, for example, a statistical question of interest concerns when a rank correlation is significantly different from 1.0. Of course, the problem of being able to reject, but not to accept, a model is a feature of classical hypothesis testing.

The results of the experiment may be summarized as follows. The model was not rejected for eight subjects, and failed for four in specific ways. Two subjects (2E and 1L) combined the diagnosticities of the events in an additive or multiplicative fashion, but violated the likelihood ratio principle. The behavior of two other subjects (6E and 4L) was completely unexplained.

However, when the model was not rejected, two interesting facts concerning the scale values emerged. The fact that $\phi_3(n_2)$ was less than 2 in all cases when $\phi_3(n_1) = 1$ was interpreted to imply that two pieces of identical evidence did not carry twice the weight of one. The relation between the ϕ_1 and ϕ_2 slopes (see Figure 2) was interpreted to indicate that the conditional probabilities of the events were differentially processed by the subjects, depending on whether they were conditional on the more or less likely spinner.

There were no apparent effects of response method on the validity of the model or on the scale values, despite the overall mean differences between the response magnitudes (cf. Table 2). It must be emphasized that the conclusions which were drawn from the experiment depended only on ordinal data. Thus, it was unnecessary to make any stronger a priori assumptions concerning the subjects' responses. It is apparent that the present approach, as illustrated in the experiment, has the advantage of testing the primary assumptions and eliminating the need for most auxiliary ones, but moreover it has provided a framework within which precise theoretical questions may arise. Thus, the focus of interest is shifted from the rather broadly formulated problems of misperception, misaggregation, and response bias, to specific ones concerning the nature and determinants of composition rules and of scale values. As is often the case, the question is not whether the model is correct in an absolute sense, but what are the conditions under which it holds (see e.g., Rapoport & Wallsten, 1972). The present experiment indicates that a more general theory is required which will predict when and how the model will fail and what relations will emerge among the scale values.

It is of interest to examine the current theoretical notions of misperception, misaggregation, and response bias in terms of the present development. Considering the levels of uniqueness of the ϕ_m, $m = 1, 2, 3$, it is impossible at this stage to arrive at "true values" of log subjective posterior odds; the scaled values will be unique up to multiplication by a positive constant. In the absence of additional theory or assumptions, questions of respones bias are meaningful only if their answers do not depend on scaling to greater degrees of uniqueness. Thus, for example, from this point of view, one may ask whether a subject's responses are nonlinearly related to his posterior odds, as one interpretation of DuCharme's (1970) position would suggest, but not whether the two are identical.

Misperception and misaggregation are not so simply treated. It is the author's feeling that the terms do not correspond precisely to anything in the present development. However, one might associate questions of aggregation with those concerning the composition rule and questions about perception with those concerning scale values. If that were done, the terms would not be equally fundamental, as they are presently in the literature, since in the spirit of conjoint measurement, scale values may be determined only after a composition rule has been validated.

Bayes' rule, as a psychological model, may be derived from Equation 5, or more generally, Equations 2b and 3b. This, in turn, is the normative model for the task under consideration. It appears that the constructs of *mis*perception and *mis*aggregation are meaningful only when the normative model is considered, or another is evaluated against it. However, as with response bias, the questions which may be asked of the data are limited by the levels of uniqueness of the scales. Therefore, one may not ask whether scale values are "correct," without additional theory or assumption.

Thus, the current problems being discussed in the literature may be viewed as a subset of those which follow from the present development. Certainly that which was presented in this paper does not constitute a complete theory of human inference on the basis of probabilistic data. However, as the present experiment demonstrates, it does provide a rigorous framework within which such a theory may unfold with a minimum number of assumptions. Work is presently planned or under way to evaluate the hypotheses which grew out of the experiment, to extend these methods to tasks in which the alternative hypotheses are verbal descriptions rather than binomial distributions, and to bring these techniques to bear in studying the processing of probabilistic information presented sequentially over time.

REFERENCES

ANDERSON, N. H. Functional measurement and psychophysical judgment. *Psychological Review*, 1970, **77**, 153–170.

BEACH, L. R. Accuracy and consistency in the revision of subjective probabilities. *IEEE Transactions on Human Factors in Electronics*, 1966, **7**, 29–37.

DUCHARME, W. M. *A review and analysis of the phenomenon of conservatism in human inference.* (Systems No. 46-5) Houston, Texas: Rice University Interdisciplinary Program in Applied Mathematics and Systems Theory, 1969.

DUCHARME, W. M. Response bias explanation of conservative human inference. *Journal of Experimental Psychology*, 1970, **85**, 66–74.

EDWARDS, W. Optimal strategies for seeking information: Models for statistics, choice reaction time, and human information processing. *Journal of Mathematical Psychology*, 1965, **2**, 312–329.

EDWARDS, W. Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment.* New York: Wiley, 1968.

EDWARDS, W., LINDMAN, H., & SAVAGE, L. J. Bayesian statistical inference for psychological research. *Psychological Review*, 1963, **70**, 193–242.

EDWARDS, W., PHILLIPS, L. D., HAYS, W. L., & GOODMAN, B. C. Probabilistic information processing systems: Design and evaluation. *IEEE Transactions on Systems Science and Cybernetics*, 1968, **4**, 248–265.

GREEN, B. F., JR. Descriptions and explanations: A comment on papers by Edwards and Hoffman. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley, 1968.

GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.

KRANTZ, D. H., LUCE, R. D., SUPPES, P., & TVERSKY, A. *Foundations of measurement*. Vol. 1. New York: Academic Press, 1971.

KRANTZ, D. H., & TVERSKY, A. Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, 1971, **78**, 151–169.

KRUSKAL, J. B., Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society*, 1965, **27**, 251–263.

LATHROP, R. G., JONES, C., & KEISER, M. Measuring decisions: A first look. Paper presented at the meeting of the Western Psychological Association, Vancouver, British Columbia, June 1969.

PETERSON, C. R., DUCHARME, W. M., & EDWARDS, W. Sampling distributions and probability revisions. *Journal of Experimental Psychology*, 1968, **76**, 236–243.

PETERSON, C. R., SCHNEIDER, R. J., & MILLER, A. J. Sample size and the revision of subjective probabilities. *Journal of Experimental Psychology*, 1965, **69**, 522–527.

PHILLIPS, L. D., & EDWARDS, W. Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 1966, **72**, 346–354.

RAPOPORT, A., & BURKHEIMER, G. J. Models for deferred decision making. *Journal of Mathematical Psychology*, 1971, **8**, 508–538.

RAPOPORT, A., & WALLSTEN, T. S. Individual decision behavior. *Annual Review of Psychology*, 1972, **23**, 131–176.

SCHUM, D. A., & MARTIN, D. W. Human processing of inconclusive evidence from multinomial probability distributions. *Organizational Behavior and Human Performance*, 1968, **3**, 353–365.

SHANTEAU, J. C. An additive model for sequential decision making. *Journal of Experimental Psychology*, 1970, **85**, 181–191.

SLOVIC, P., & LICHTENSTEIN, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. In L. Rapoport & D. A. Summers (Eds.), *Human judgment and social interaction*. New York: Holt, Rinehart & Winston, 1972, in press.

SWETS, J. A., & BIRDSALL, T. G. Deferred decision in human signal detection: A preliminary experiment. *Perception and Psychophysics*, 1967, **2**, 15–28.

ULEHLA, Z. J., CANGES, L., & WACHWITZ, F. Integration of conceptual information. *Psychonomic Science*, 1967, **8**, 223–224.

ULEHLA, Z. J., HALPERN, J., & CERF, A. Integration of information in a visual discrimination task. *Perception and Psychophysics*, 1968, **4**, 1–5.

VLEK, C. A. J., & VAN DER HEIJDEN, L. H. C. *Subjective likelihood functions and variations in the accuracy of probabilistic information processing.* (Rep. No. E017-67) The Netherlands: Psychological Institute, University of Leiden, 1967.

WALLSTEN, T. S. Subjectively expected utility theory and subjects' probability estimates: Use of measurement-free techniques. *Journal of Experimental Psychology*, 1971, **88**, 31–40.

WHEELER, G., & BEACH, L. R. Subjective sampling distributions and conservatism. *Organizational Behavior and Human Performance*, 1968, **3**, 36–46.

ZINNES, J. L. Scaling. *Annual Review of Psychology*, 1969, **20**, 447–478.

APPENDIX

Limiting consideration to samples containing n_k replicates of E, using the notation n_kE instead of D_n, and making Assumption *e*, that the effect of position on an event's diagnosticity is independent of the $P(E|X_i)$ and $P(E|Y_j)$ occurring at that position and identical on both dimensions, the following theorem is proved:

Theorem

The model

$$R(X_i, Y_j | n_k E) = g\{\phi_3(n_k)[\phi_1(E | X_i) - \phi_2(E | Y_j)]\}$$

holds if and only if

(*a*) $R(X_i, Y_j | n_k E) = g[\phi_1(n_k E | X_i) - \phi_2(n_k E | Y_j)]$,

(*b*) $R(X_i, Y_j | n_k E) = g[\psi_1(E) + \cdots + \psi_{n_k}(E)]$,

(*c*) Assumption *e* holds.

Proof

It is shown first that the model follows from *a*, *b*, and *c*. Applying *a* and *b* when only the first event of the sample has occurred yields

$$\psi_1(E) = \phi_{11}(E | X_i) - \phi_{21}(E | Y_j), \quad [A1]$$

where the second subscript on both ϕ_m denotes sequence position number, h.

Applying *a* and *b* again for a second event in the sample, that is, when $n_k = 2$, yields

$$\psi_1(E) + \psi_2(E) = \phi_1(2E | X_i) - \phi_2(2E | Y_j).$$

Substituting from Equation A1 and rearranging terms:

$$\psi_2(E) = [\phi_1(2E | X_i) - \phi_{11}(E | X_i)] - [\phi_2(2E | Y_j) - \phi_{21}(E | Y_j)].$$

The terms within the brackets may be considered as $\phi_{12}(E | X_i)$ and $\phi_{22}(E | Y_j)$, respectively. Combining this with Equation A1 yields

$$\psi_1(E) + \psi_2(E) = \phi_{11}(E | X_i) + \phi_{12}(E | X_i) - \phi_{21}(E | Y_j) - \phi_{22}(E | Y_j). \quad [A2]$$

It is evident from the development leading to Equation A2, that *a* and *b* imply the general additive model for all n_k.

$$R(X_i, Y_j | n_k E) = g[\phi_{11}(E | X_i) + \cdots + \phi_{1n_k}(E | X_i) - \phi_{21}(E | Y_j) - \cdots - \phi_{2n_k}(E | Y_j)]. \quad [A3]$$

Without loss of generality, let $\phi_{1h}(E | X_i) = \alpha_h \phi_{11}(E | X_i)$ and $\phi_{2h}(E | Y_j) = \beta_h \phi_{21}(E | Y_j)$, for all h, with $\alpha_1 = \beta_1 = 1$. Invoking *c*, $\alpha_h = \beta_h$, for $h = 1, \cdots, n_k$. Substituting into Equation A3 and rearranging,

$$R(X_i, Y_j | n_k E) = g\{[1 + \alpha_2 + \cdots + \alpha_{n_k}] \times [\phi_{11}(E | X_i) - \phi_{21}(E | Y_j)]. \quad [A4]$$

The distributive model follows by letting $\phi_3(n_k) = 1 + \alpha_1 + \cdots + \alpha_{n_k}$ in Equation A4 and by dropping the subscripts for sequence position, h, from the ϕ_{mh}. The proof is completed by working through the steps in reverse order to show that *a*, *b*, and *c* are implied by the model.

(Received November 11, 1971)

Psychological Review
1972, Vol. 79, No. 3, 261–264

THEORETICAL NOTES

ON TESTING THE INDEPENDENCE OF ASSOCIATIONS

DOUGLAS L. HINTZMAN[1]

University of Oregon

The failure of chi-square independence tests to demonstrate a significant relationship between retrieval of A–B and A–C associations in modified-modified free recall does not imply that the associations are independent. The negative correlation between B and C recall predicted by interference theory may be masked by a positive correlation due to differences among subjects and items, or to mediation. Evidence for the existence of these confounding factors in A–B, A–C experiments is reviewed. A correlational analysis such as the chi-square test is not a satisfactory method of establishing the independence (or dependence) of associations.

Greeno (1969) and Martin (1971) have drawn some rather strong theoretical conclusions from what Martin calls the "DaPolito independent retrieval phenomenon." The phenomenon is as follows: Following the learning of two paired-associate lists having an A–B, A–C relationship, subjects are given a modified-modified free recall (MMFR) test—an unpaced recall test in which each A item is presented and the subject attempts to respond with both the B and C responses. Each subject-item is then tallied in a 2 × 2 contingency table, according to whether the first-list response was recalled (B vs. $\bar{\text{B}}$) and whether the second-list response was recalled (C vs. $\bar{\text{C}}$). Chi-square tests on frequencies obtained in this manner in a number of experiments are remarkably consistent in failing to show a significant relationship between B and C recall (DaPolito, 1966; Greeno, 1969; Martin, 1971). It has been concluded that the existence of the A–C association does not affect retrieval of A–B, and vice versa; that is, that the associations are independent. The conclusion is important because it contradicts two hypotheses that have been made in the context of interference theory: (*a*) that A–B associations are unlearned during A–C acquisition, and (*b*) that specific response competition affects MMFR performance. The conclusion that associations are independent has been stated especially emphatically by Martin (1971). However, this paper will show that the conclusion does not follow from the chi-square independence tests presented in its support.

The interference hypothesis appears to predict a negative relationship between recall of B and C. That is, that $P(\text{B}|\text{C}) < P(\text{B})$ and $P(\text{C}|\text{B}) < P(\text{C})$, or stated in terms of entries in the 2 × 2 table, that the B$\bar{\text{C}}$ and $\bar{\text{B}}$C cell frequencies should be higher than expected by "chance." This is the dependence that the chi-square independence tests have failed to uncover. This evidence, however, is correlational, and a zero correlation no more implies the absence of a causal relationship than a nonzero correlation implies its presence. In fact it is quite possible in this case that the expected negative correlation is being masked by factors which tend to produce a positive correlation between B and C recall.

The most serious problem is that the procedure just outlined pools data over subjects and stimulus terms. Thus, quite apart from any theoretical prediction, one should expect to find a positive relationship between recall of B and C. Good subjects will tend to recall both B and C while poor subjects will recall neither, and easily encoded stimuli will tend to elicit both responses while hard to encode stimuli will not. Other factors being absent, then, subject and item differences dictate that the BC and $\overline{\text{BC}}$ cells of the table should contain the high frequencies.

The fact that subject and item differences can be expected to produce an effect opposed to that of interference appears to have escaped Martin's (1971) notice. He deals with subject differences by saying:

> Strictly speaking, such chi-squares may be illegitimate, inasmuch as for a given 2 × 2 table each subject contributed to more than one cell. Thus, the independent-observations assumption that underlies

---

[1] Requests for reprints should be sent to Douglas L. Hintzman, Department of Psychology, University of Oregon, Eugene, Oregon 97403.

TABLE 1

MMFR Performance of Hypothetical Subjects of High, Medium, and Low Ability

Response	High[a]		Medium[b]		Low[c]	
	C	$\bar{C}$	C	$\bar{C}$	C	$\bar{C}$
B	6	2	2	3	0	2
$\bar{B}$	2	0	3	2	2	6

[a] $N = 25$; $\chi^2 = 15.6$.
[b] $N = 50$; $\chi^2 = 20.0$.
[c] $N = 25$; $\chi^2 = 15.6$.

> the chi-square test was violated to the extent that there was an interaction between subjects and items in the experiment. But such a violation can only inflate chi-square [p. 318].

The final sentence is untrue. This will be illustrated with an example involving subject differences; however the problem is much more general than a simple violation of the independent-observations assumption. It would be present even if each subject contributed only one observation to the table and even if all subjects were essentially the same, since there are other likely sources of positive correlation.

For convenience, the example assumes that lists are 10 items long and that subjects fall into three distinct ability categories: high, medium, and low. Table 1 illustrates the MMFR performance of each hypothetical type of subject. Note that the performance of each type of subject conforms to the interference hypothesis, in that cells $B\bar{C}$ and $\bar{B}C$ contain frequencies higher than expected by chance. Now, assume that there are 100 subjects in all: 25 of high, 50 of medium, and 25 of low ability. Taken separately, chi-squares computed on frequencies of each of these three groups are highly significant (15.6, 20.0, and 15.6, respectively), reflecting the effect of "interference." However, if the response of all 100 subjects are now included in a single 2 × 2 table, each cell entry will be exactly 250, and chi-square will be zero. The positive correlation between B and C recall induced by subject differences completely masks the negative correlation due to interference. Notice that this effect is not simply due to the fact that each subject contributed more than one observation to the table. If we were to assume only one observation per subject and multiply the number of subjects of each type by 10, all the frequencies and the obtained chi-squares would be unchanged.

One might object that the ability differences assumed in Table 1 are extreme. But they serve to illustrate the general point, and the present argument does not rest on subject differences alone. Differences in item difficulty could be substituted for subject differences in the table with the same result. In general, any source of positive correlation between recall of B and C will tend to counteract the effect of interference. A third source, in addition to subject and item differences, involves the use of List 1 associations to mediate List 2 learning.

The existence of these confounding factors is not just conjecture. There is evidence of their presence in several A–B, A–C experiments, including some that Greeno (1969) and Martin (1971) have cited as providing evidence for independent retrieval.

Subject Differences

Differences among subjects within a group are sometimes assumed to be inconsequential because a single estimate of a parameter value in a stochastic model enables the model to fit the variance, as well as the mean of the entire group. A more analytic test for subject differences involves the relationship between learning rates on two successive lists. In an unpublished A–B, A–C transfer experiment, the present author obtained rank-order correlations of .93, .30, .66, and .60 between List 1 and List 2 learning rates in four different conditions. Postman and Stark (1965) reported a like correlation of .45. They also obtained partial correlations of .55 and .75 between acquistion of List 2 and MMFR performance on responses from List 1. Since List 2 acquisition and List 2 MMFR are likely to be directly related, this can be taken as evidence that subject differences contribute substantially to a positive relationship between B and C on the MMFR test. Further, it cannot be reasonably claimed that subject differences are minimized by the acquisition criterion the experimenter adopts. For one thing, the studies cited by Martin (1971) as showing independence were not consistent in their use of a practice or a performance criterion, and several even used one criterion for List 1 and another for List 2. For another, even if the learning criterion were standardized, there would be no guarantee that "degree of learning" had really been equated (cf. Underwood, 1964). Indeed, the relationship between learning criterion and recall appears to be quite complex. Postman and Stark (1965) report that when List 1 is learned to mastery, slow learners recall the B items better than fast learners—the opposite of what would be obtained if a fixed number of trials were given.

Item Differences

Since in an A–B, A–C paradigm the stimulus terms in the two lists are identical, the rates at which corresponding associations are learned should be correlated simply because of differences in the ease with which the stimuli can be reliably encoded. Meaningfulness, image-evoking power, and similarity to other stimuli in the list are all possible factors here, and even though an attempt is usually made to equate the stimuli on these dimensions, subject-item interactions are still bound to be present. In an early study cited as showing evidence for independent retrieval, Runquist (1957) noted that learning rates of corresponding A–B and A–C pairs were correlated, but dismissed the correlation as a factor in recall. Positive relationships of the same kind have been reported by Dallett and D'Andrea (1965), Postman (1964), Postman and Stark (1969), and Wichawut and Martin (1971). The Wichawut and Martin data suggest, in addition, that the correlation persists in MMFR. Their Figure 3 shows that when A–B pairs were rank ordered according to difficulty in learning, MMFR performance on both B and C responses was directly related to the A–B ranking. Interestingly, when item differences were taken into consideration in this way, some evidence for interference emerged: under low degrees of learning, the relationships $P(\mathrm{B}|\mathrm{C}) < P(\mathrm{B})$ and $P(\mathrm{C}|\mathrm{B}) < P(\mathrm{C})$ were observed. This happened despite the fact that the conditional probabilities $P(\mathrm{B}|\mathrm{C})$ and $P(\mathrm{C}|\mathrm{B})$ tend to select the better subjects, and so might well be expected to exceed the unconditional values if, for a given subject, the two associations were truly independent.

Mediation

A third potential source of correlation between B and C recall in MMFR involves attempts by subjects to mediate List 2 learning using associations acquired in List 1. To the extent that these attempts are successful, they can be expected to affect a positive relationship. Evidence regarding mediation in the A–B, A–C paradigm has been presented by Dallett and D'Andrea (1965), Martin and Dean (1966), and Palermo (1962).

An obvious question, in light of these reasons for expecting a positive correlation between B and C performance, is why the chi-square tests have been in such consistent agreement with the null hypothesis of independence. It is tempting to conclude that the positive relationship is being masked by a negative one, as would be produced by interference between the A–B and A–C associations. Thus, one could turn Martin's (1971) argument around and claim that the "independent retrieval phenomenon" is evidence for, rather than against, interference. This argument, however, rests on the same correlational evidence as Martin's and is suspect for the same reasons. Only the most tentative conclusion can be drawn from an analysis that fails to take into account such gross sources of selective bias as variability among subjects and items.

Is the retrievability of association A–B independent of the existence of A–C? The best way to answer this question is to experimentally manipulate the existence of A–C, rather than trying to infer its existence from the subject's performance. In the Wichawut and Martin (1971) experiment, 25% of the List 1 stimuli were not used in List 2 pairs, and so their A–B associations should have suffered no specific interference from List 2 learning. The B responses to those stimuli were better recalled in MMFR than the B responses to stimuli which did occur in List 2. And although her results failed to reach significance, Birnbaum (1970) obtained differences in the same direction. This outcome is clearly consistent with the interference hypothesis and lends support to the present arguments concerning the inadequacy of chi-square tests as measures of the independence of associations.

REFERENCES

BIRNBAUM, I. M. Response selection in retroactive inhibition. *Journal of Experimental Psychology*, 1970, **85**, 406–410.

DALLETT, K. M., & D'ANDREA, L. Mediation instructions versus unlearning instructions in the A–B, A–C paradigm. *Journal of Experimental Psychology*, 1965, **69**, 460–466.

DAPOLITO, F. J. Proactive effects with independent retrieval of competing responses. Unpublished doctoral dissertation, Indiana University, 1966.

GREENO, J. G. *A cognitive interpretation of negative transfer and forgetting in paired-associates.* (Human Performance Center Memorandum Rep. No. 9) Ann Arbor: University of Michigan, 1969.

MARTIN, E. Verbal learning theory and independent retrieval phenomena. *Psychological Review*, 1971, **78**, 314–332.

MARTIN, R. B., & DEAN, S. J. Reported mediation in paried-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 23–27.

PALERMO, D. S. Mediated association in a paired-associate transfer task. *Journal of Experimental Psychology*, 1962, **64**, 234–238.

POSTMAN, L. Studies of learning to learn: II. Changes in transfer as a function of practice. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 437–447.

POSTMAN, L., & STARK, K. The role of response set in tests of unlearning. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 315–322.

POSTMAN, L., & STARK, K. The role of associative mediation in retroactive inhibition and facilitation. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 790–798.

RUNQUIST, W. N. Retention of verbal associates as a function of strength. *Journal of Experimental Psychology*, 1957, **54**, 369–375.

UNDERWOOD, B. J. Degree of learning and the measurement of foregetting. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 112–129.

WICHAWUT, C., & MARTIN, E. Independence of A–B and A–C associations in retroaction. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 316–321.

(Received August 31, 1971)

Psychological Review
1972, Vol. 79, No. 3, 265-267

INDEPENDENCE OF ASSOCIATIONS TESTED:

A REPLY TO D. L. HINTZMAN

EDWIN MARTIN[1] AND JAMES G. GREENO

University of Michigan

D. L. Hintzman's reservation is well taken. The prescribed more detailed analyses of possible independence between A–B and A–C associations reveal, however, that independence holds for good and poor subjects, for easy and difficult items, and for any combination of subjects and items.

Hintzman's (1972) point is that differences among subjects and among items may systematically combine with associative interference effects to yield the DaPolito independent retrieval phenomenon—a nonsignificant chi-square when testing for independence of A–B and A–C associations in a subsequent stimulated recall of B and C responses. Poor subjects and difficult items may produce a high incidence of recalling neither B nor C; good subjects and easy items may produce a high incidence of recalling both B and C. These two effects together could produce a joint incidence of both-B-and-C and neither-B-nor-C outcomes that combines with the associative interference prediction of either B or C but not both to produce a picture of independence that is spurious. We accept this point as an important and critical possibility. Our principal response is based on empirical data.

Consider first the experiment by Wichawut and Martin (1971). After A–B learning, the degree of A–C learning was manipulated within subjects. Then a standard test for retroaction was administered in which the A stimuli were presented singly, and the subject was requested to respond to each with both B and C as best he could (the modified-modified free recall [MMFR] test). Of the 36 subjects, 8 had 100% recall of the C responses and hence were dropped from this analysis. For each of the remaining 28 subjects, unconditional $P(B)$ and $P(C)$ and joint $P(B \& C)$ were computed from the MMFR results. In Figure 1, $P(B \& C)$ is plotted against the product $P(B)P(C)$. The dashed contour marks binomial standard deviations based on $P(B)P(C)$. It is our opinion that the observed deviations from the identity diagonal are not systematic over subjects. Independence holds for poor (lower left-hand quadrant) as well as good (upper right-hand quadrant) subjects.

The number of pairs in the lists used by Wichawut and Martin is 16. Using stimuli as units, we carried out the same independence analysis. The result is shown in Figure 2. As with subjects in the preceding paragraph, we conclude the following for independence: There is no systematic deviation from the identity diagonal over items.

Of the 36 subjects, we selected the best 9 and the worst 9 on the basis of first-task (A–B) performance. Independently, we then selected the four best and four worst stimuli, again on the basis of first-task performance. The 18 remaining subjects and 8 remaining stimuli were labeled "middle." For each of the nine combinations of best, middle, and worst stimuli and subjects, we calculated product $P(B)P(C)$ and joint $P(B \& C)$ from the MMFR results. These are posted in Table 1. Each observed $P(B \& C)$ is well within the binomial standard deviation of predicted $P(B)P(C)$.

Note that the marginal results for subjects in Table 1 (right-hand column) correspond to Hintzman's Table 1. Observed $P(B \& C)$ clearly varies according to whether the subjects were best, middle, or worst. But this variation is exactly as predicted by the independence hypothesis, $P(B)P(C)$.

Other researchers, without prompting, have been sensitive to the possibility raised by Hintzman. Malis (1970) conducted two A–B, A–C experiments. For each experiment he calculated from MMFR results individual-subject chi-squares to test the independence of A–B and A–C. To so proceed is to represent each subject's deviation from independence by a chi-square with one degree of freedom. Summing n such chi-squares gives a chi-square with n degrees of freedom. This is a chi-square that can be large to the extent that good subjects get both B and C and that poor

---

[1] Requests for reprints should be sent to Edwin Martin, Human Performance Center, 330 Packard Road, Ann Arbor, Michigan 48104.

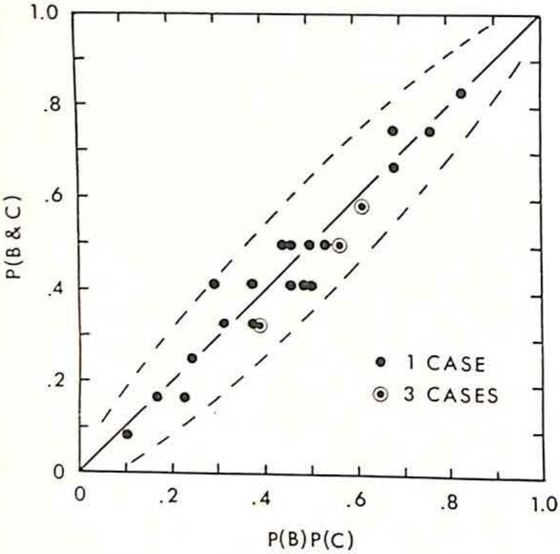


FIG. 1. Observed joint probability of B and C against expected joint probability for 28 individual subjects. (Dashed contour is binomial standard deviation of $P(\text{B})P(\text{C})$.)

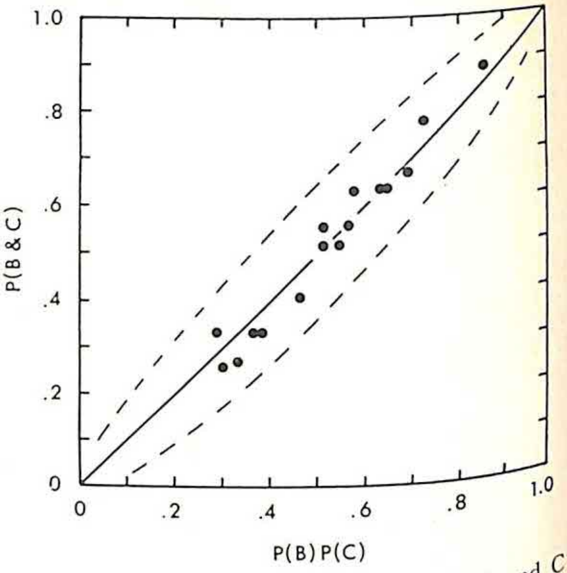


FIG. 2. Observed joint probability of B and C against expected joint probability for 16 different stimuli. (Dashed contour is binomial standard deviation of $P(\text{B})P(\text{C})$.)

subjects get neither B nor C; that is, possible differential directional deviations from independence of good and poor subjects cannot cancel each other. Malis reports nonsignificant chi-squares throughout.

In the first study to call attention to statistical independence of responses in the MMFR test, DaPolito (1966, Experiment I) tested for independence within the best, intermediate, and worst subjects in his experiment. The results of his analysis are like those in the right-hand column of Table 1 herein, and accordingly unlike the hypothetical subjects described in Hintzman's Table 1. The largest of the three chi-squares reported by DaPolito is 1.07 ($p = .30$).

TABLE 1

PREDICTED AND OBSERVED JOINT OCCURRENCES OF B AND C RESPONSES FOR BEST, MIDDLE, AND WORST SUBJECTS AND STIMULI

Subjects	Statistic	Stimuli			Marginal
		4 best	8 middle	4 worst	
9 best	$P(\text{B})P(\text{C})$	.773	.690	.462	.658
	$P(\text{B \& C})$	.793	.686	.462	.660
18 middle	$P(\text{B})P(\text{C})$	.717	.501	.408	.532
	$P(\text{B \& C})$	.715	.472	.389	.514
9 worst	$P(\text{B})P(\text{C})$	.386	.343	.408	.369
	$P(\text{B \& C})$	.391	.368	.357	.370
Marginal	$P(\text{B})P(\text{C})$	.655	.499	.430	.520
	$P(\text{B \& C})$	.667	.495	.398	.514

In addition to these empirical results, a methodological point can be made regarding Hintzman's argument. He remarks that failure to reject the null hypothesis of independence in data "does not imply that the associations are independent." He is certainly correct. Neither do the results given in this rejoinder imply that associations are independent. Fact (or existential) statements do not imply hypothetical (or universal) statements. The relation of implication goes in the opposite direction: Hypotheses imply facts. Strong theoretical hypotheses imply assertions that are testable in the sense that falsifying observations are in principle possible.

The hypothesis of A–B, A–C independence is that the strength or effectiveness of one association involving A does not affect the strength or effectiveness of the other, and that emitting B and C responses does not produce a competition effect on performance in MMFR tests. This is a testable hypothesis because it implies that responses in an MMFR test will be statistically independent. Observation of statistical dependence between B and C responses is all that is required to deny the independence hypothesis. Extensive data have failed to deny this hypothesis (Greeno, 1969; Martin, 1971), and so have the stronger tests (this rejoinder) that Hintzman thought might lead to denial. There are thus considerable grounds (inductive, not implicative) for acceptance of the idea that independence holds for individual subjects, individual items, and subject-item combinations.

In contrast, associative interference theory, when coupled with reasonable considerations of subject and item effects, allows for any relation between B and C responses that might occur in data, a point nicely made by Hintzman whether intentionally or not. Accordingly, this form of associative interference theory is essentially untestable; it has far less empirical content than is acceptable or warranted by present knowledge.

In summary, we have good reasons for accepting the independence hypothesis, and no reason to reject it. As for the associative interference hypothesis: *Without* consideration of possible subject and item effects, we have every reason to reject it, namely, the DaPolito independent retrieval phenomenon; *with* allowance for possible subject and item effects, it is too open-ended to test and hence of little, if any, theoretical interest.

REFERENCES

DAPOLITO, F. J. Proactive effects with independent retrieval of competing responses. Unpublished doctoral dissertation, Indiana University, 1966.

GREENO, J. G. *A cognitive interpretation of negative transfer and forgetting in paired-associates.* (Human Performance Center Memorandum Rep. No. 9) Ann Arbor: University of Michigan, 1969.

HINTZMAN, D. L. On testing the independence of associations. *Psychological Review,* 1972, **79,** 261–264.

MALIS, G. The priority effect in the A–B, A–D paradigm. Unpublished doctoral dissertation, University of Toronto, 1970.

MARTIN, E. Verbal learning theory and independent retrieval phenomena. *Psychological Review,* 1971, **78,** 314–332.

WICHAWUT, C., & MARTIN, E. Independence of A–B and A–C associations in retroaction. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10,** 316–231.

(Received November 10, 1971)

Psychological Review
1972, Vol. 79, No. 3, 268–274

A STOCHASTIC MODEL FOR FREE ASSOCIATION RESPONSE HIERARCHIES?[1]

VAIRA VĪĶIS-FREIBERGS[2]

Université de Montréal

A stochastic model for a class of skew distribution functions has been developed by H. A. Simon in 1955 and applied to free association data by W. J. Horvath in 1963. This paper shows that a basic assumption of the stochastic model is not satisfied by free association data and that consequently the stochastic model is not applicable in the sense proposed by Horvath. Horvath's claim that α, a parameter of the stochastic model, allows for a simple comparison of results obtained under different empirical situations is shown to be seriously misleading, since α varies systematically with sample size for both free association and continuous prose data (where it is known as the type–token ratio). Finally, Horvath's assertion that α is a better measure of associative "power" than the frequency of the primary response is also questioned (since each measures quite different aspects of free association response hierarchies), and a systematic comparison of the properties of α and the frequency of the primary response is presented.

Horvath (1963) has described a new statistic, α, which is said to provide a convenient quantitative measure for comparing free association hierarchies obtained under different empirical situations. According to Horvath, α is a better measure of the association power of a word than is the frequency of the primary response (R_1).

The derivation of α presented by Horvath is based on Simon's (1955) development of a general probability function for a particular class of skewed distributions. These distributions, which appear in a wide range of empirical data from various fields, can be obtained as stationary solutions of certain stochastic processes. Only two relatively weak probability assumptions are needed for the derivation of the stochastic model.

Horvath's presentation assumes that both these basic assumptions are satisfied by free association data. The aim of the present paper is to show that this is not the case and to point out the serious difficulties that might arise in an uncritical application of the general stochastic model. Comparisons of α and R_1 will also be made in order to clarify the nature of the information conveyed by each of these measures. First of all, however, a brief resume of the general case of Simon's stochastic model is necessary.

Basic Stochastic Model

The general case of Simon's stochastic model is described in terms of word frequencies in prose samples. Consider a sample from a continuous source of words (e.g., a book) that has reached a length of k words. The number of different words having occurred exactly i times in these first k words is $f(i, k)$. The basic probability assumptions of the model can be stated as follows:

Assumption 1: The probability that the $(k + 1)$*st word is a word that has already appeared exactly* i *times is proportional to* $i \times f(i, k)$. That is, it is proportional to the total number of occurrences of all the words that have appeared exactly i times (Simon, 1955, p. 427). This allows for the possibility that among a group of words with the same frequency i within a sample k, the probability of recurrence of some may be much higher than that of others.

Assumption 2: There is a constant probability, α*, that the* $(k + 1)$*st word be a new word—a word that has not occurred in the first* k *words.* This can be expressed as $\alpha = \Delta n_k / \Delta k =$ a constant, where n_k is the number of different words, k is the total number of word occurrences, and α is the rate at which new words are introduced.

As can be seen from the formula, α is equivalent to the type–token ratio, a measure of vocabulary diversity long familiar in statistical investigations of language (cf. Estoup, 1917).

Since, according to the stochastic model, all frequencies grow proportionately with sample size, k (Assumption 1), the relative frequencies, f^*i (i.e., i/k), are independent of k. Similarly α, the rate at which new words are introduced, is also independent of k for the general case (Assumption 2).

---

[1] This work was supported by Grant APA-219 from the National Research Council of Canada.

[2] Requests for reprints should be sent to Vaira Viķis-Freibergs, Institut de Psychologie, Université de Montréal, 90, rue Bellingham, Montréal 101, Qué., Canada.

The general case, however, only applies where new words are introduced at a constant rate. We shall now show that in free association data, where the first assumption seems reasonably well satisfied, the second assumption is not tenable.

Relation of α to Sample Size: Empirical Data

Test of basic Assumption 2. If Assumption 2 of the basic stochastic model is valid, the values of α should remain comparable for association norms based on widely differing sample sizes. Such a comparison can be conveniently made with the Connecticut and the Minnesota free association norms, which use samples of 150 and 1,008 subjects, respectively, and have 62 stimulus words in common (Bousfield, Cohen, Whitmarsh, & Kincaid, 1961; Russell & Jenkins, 1954).

The effect of sample size on the spread of the response distribution can be seen in Table 1 for a set of five different stimulus words. Table 1 shows that, as k increases by 858 between the two samples, the increments in n_k are only in the range from 16 to 56. It is evident that the increase in new words (k) is disproportionately larger than the increase in different words (n_k). As a result, α is systematically much smaller for the larger sample (for the word TABLE, e.g., α is nearly five times less for the Minnesota than for the Connecticut data).

That these are not isolated examples can be seen by comparing α values for the entire set of 62 Kent-Rosanoff words which are common to the Minnesota and the Connecticut free association norms. The mean α is .236 for the Connecticut data ($k = 150$) and .100 for the Minnesota data ($k = 1{,}008$).[3] The difference between these means is highly significant ($t = 15.28$, $df = 122$, $p < .001$). There is also a change in the shape of the distribution of α, which is normal for the large sample, but tends to be rectangular for the Connecticut data.

These results, showing significant changes in α with increasing sample size, are incompatible with Assumption 2 of the stochastic model.

TABLE 1

EFFECT OF SAMPLE SIZE ON NUMBER OF DIFFERENT RESPONSES (n_k) AND ON α

Stimulus word	Connecticut norms $k = 150$		Minnesota norms $k = 1008$	
	n_k	α	n_k	α
TABLE	31	.207	47	.046
BLACK	32	.213	78	.077
SHEEP	46	.307	101	.100
PRIEST	42	.280	98	.097
CITY	45	.300	92	.091

Since Assumptions 1 and 2 are not only sufficient but also necessary in Simon's basic stochastic model, a direct application of this model to word association data, as done by Horvath (1963), is not justifiable.

Test of modified Assumption 2. While the basic stochastic model, which is the one Horvath used, assumes that α is independent of k, Simon does present two generalizations allowing for modification of this assumption. The second one of these assumes a case where α decreases with k, but very slowly (Simon, 1955, "Case II," p. 431). Thus, should the departure of the main parameter from stationariness not be too pronounced, appropriate substitutions in the equation giving expected frequencies could be made, and the advantages of having a mathematical model could be preserved. The question we now have to test is whether this modified Assumption 2 is satisfied by word association data.

In order to evaluate the rate of change of α as a function of k, sample size was systematically varied for the same set of stimulus words. The free association data used were the Montreal French-Canadian norms for the 100 Kent-Rosanoff stimuli (Freibergs, 1968). The maximum number of subjects was 300, the sample size for the published norms. A computer program was used[4] which takes successively larger samples from this pool of subjects and prints out the response hierarchies and the values of α at each level. Sampling at each level of k was done without replacement, since each subject should not contribute more than one response to the total data pool.

---

[3] For those words to which some of the subjects failed to respond, the values of n_k and k were adjusted accordingly.

[4] Grateful thanks for writing the program are due to I. F. Freibergs, Département d'Informatique, Université de Montréal.

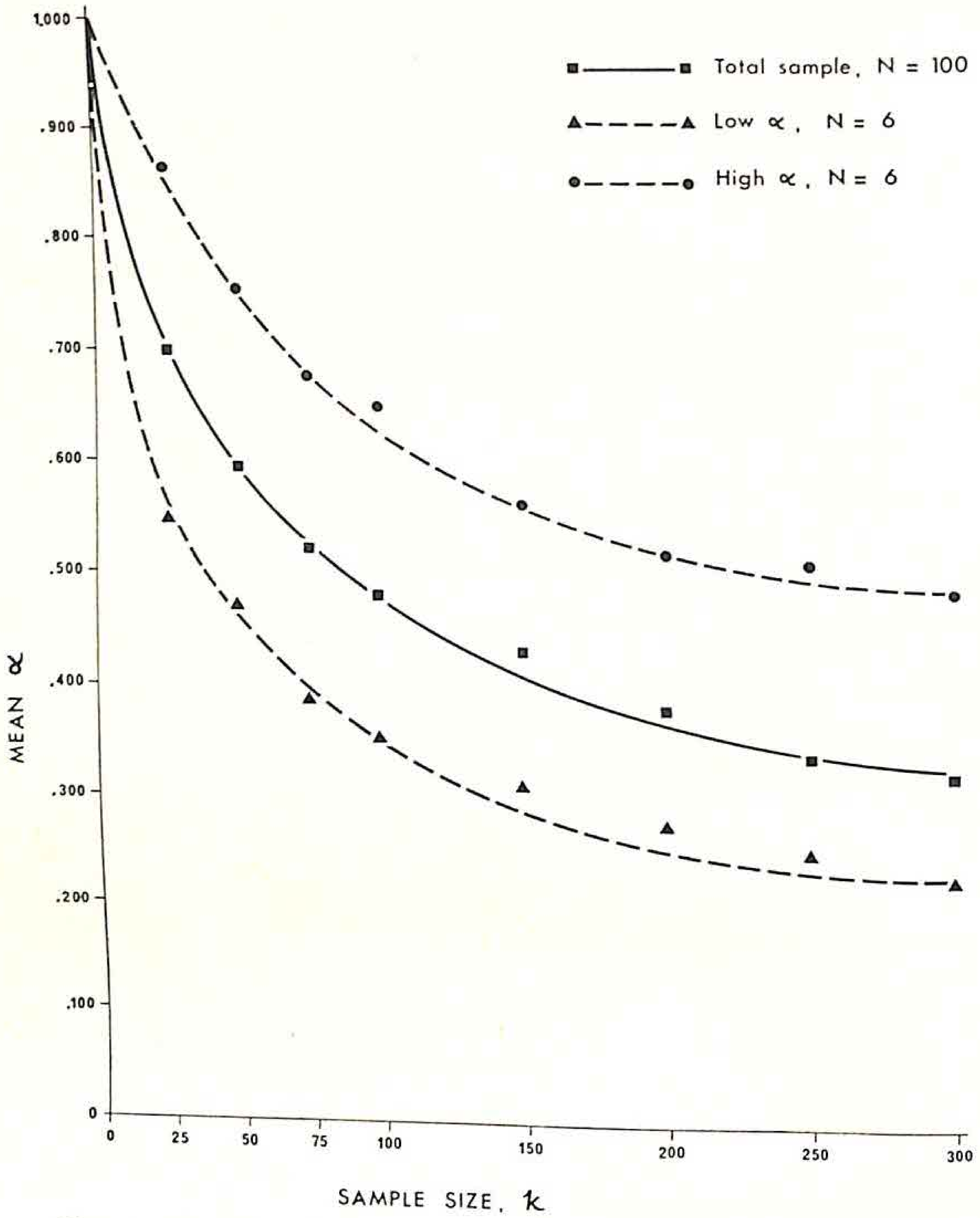


FIG. 1. Variations in α with sample size, k, for free association data.

The central curve in Figure 1 shows the mean values of α (averaged over response distributions to the 100 Kent-Rosanoff stimuli) for successively larger sample sizes. It can be seen that α is essentially a monotonically decreasing function of k. However, instead of a slow and gradual decrease in the values of α, as required by the stochastic model, there is a very rapid decrease between $k = 1$ and $k = 100$, followed by a leveling off between $k = 150$ and $k = 300$.

Separate curves were also drawn for means of the six response distributions having the highest and the lowest α values, respectively. As can be seen in Figure 1, the shape of these curves remains remarkably similar to the average curve.

Polynomial regressions of the first and second order were fitted to the mean α curve, as well as to the 100 separate curves for each different Kent-Rosanoff stimulus. While all curves gave a reasonably good fit with a first-order (linear) equation, significant quadratic components were also obtained for all curves, and the fit of the data with a second-order (quadratic) equation was always improved. Correlations with the theoretical points of the quadratic equations were highly significant in all cases (for 97 of the 100 curves, $p < .001$, 91 of these having numerical values between .90 and .99).

The first-order equation for the average α curve was $y = .755 - .174x$, which gave a correlation coefficient of .85 with the data.

The second-order equation for the average α curve was $y = .882 - .489x + .106x^2$, which gave a correlation coefficient of .95 with the data.

Implications of the Empirical Findings

Horvath has claimed three main advantages to applying Simon's basic stochastic model to free association data: (*a*) the advantages of having a mathematical model, mainly the possibility of calculating expected frequencies; (*b*) the advantages of having a new descriptive statistic, α; and (*c*) the advantages of the convenience of this statistic for comparing data from different sources. These claims will be examined in turn on the basis of the empirical evidence presented earlier.

a. First, the systematic changes in α with k, by violating Assumption 2 of the model, also violate the basic requirement of a "steady-state" distribution for calculating expected frequencies. If decreases in α with k were more gradual, then the approximate solution proposed by Simon could be used. But, as was shown, the decrease in α is on the contrary quite abrupt. As a result, the stochastic model adds nothing substantial to Skinner's earlier (1937) invocation of Zipf's law to association data, the distribution function presented by Horvath being equivalent to Zipf's law.

One objection which might be raised to the above argument is that the function of α with k might tend to reach an asymptote, and that expected frequencies could be calculated for large values of k. This could only be verified empirically by extending α curves for free association data to cover substantially larger samples. The equivalent data are available, however, for continuous prose samples, and Figure 2 shows two such functions of α over k, one for a sample of 3,000 (Chotlos, 1944) and another for a sample of 1,000,000 words (Kučera & Francis, 1967). These curves can be seen to be quite similar to the α curves for free association shown in Figure 1, and it would appear that they do approach, but by no means reach, an asymptote with increasing k.

It must also be pointed out that even with approximations for calculating expected frequencies, these could only be calculated for the tail end of the distribution (i.e., giving the expected numbers of subjects producing the most original responses). The frequencies of the most popular responses cannot be determined on the basis of the stochastic model.

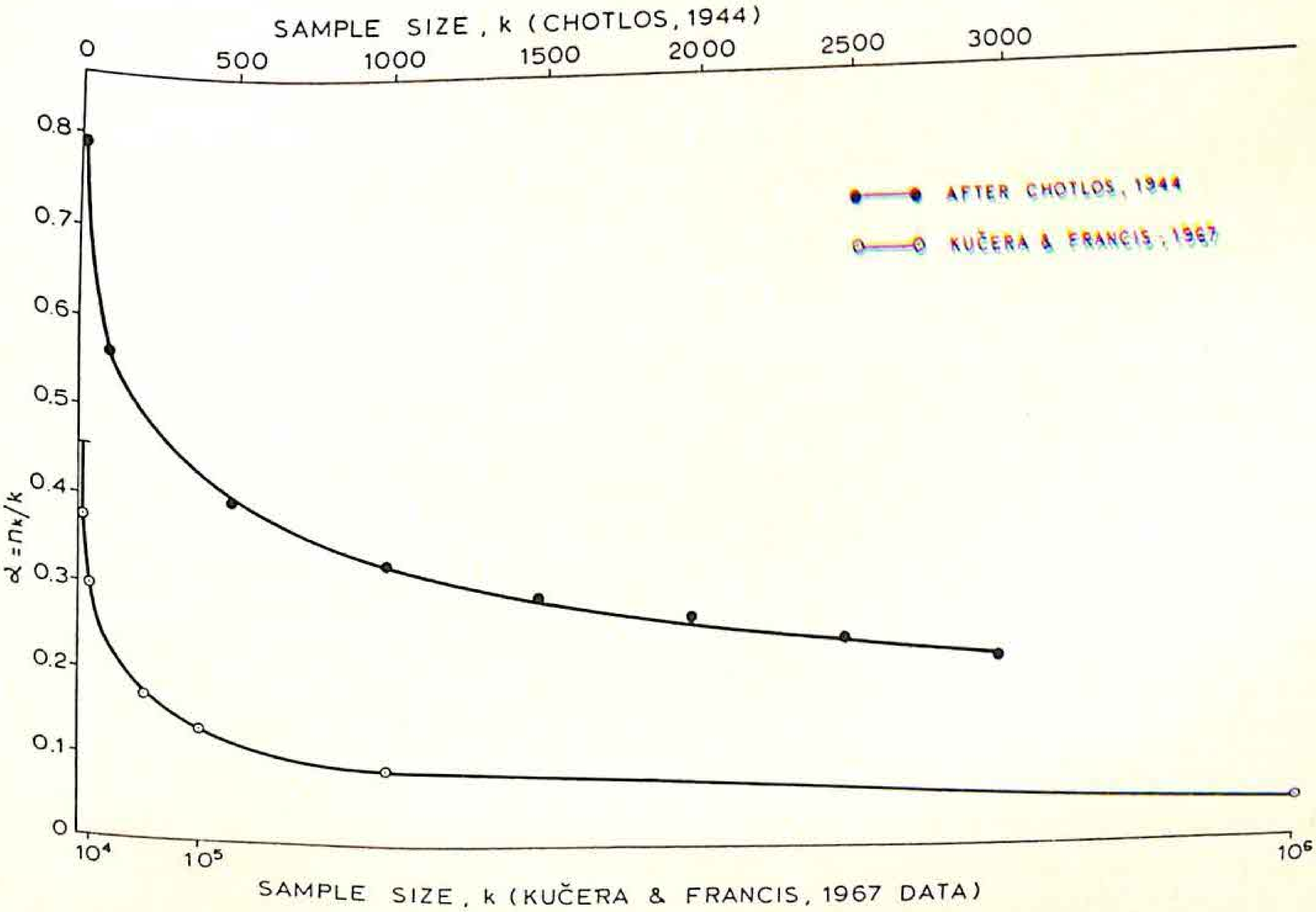


FIG. 2. Variations in α or type–token ratio with sample size, k, for continuous prose data.

b. As indicated earlier, α is not precisely new as a descriptive statistic of verbal diversification, since it goes back, under the name of type–token ratio or some other, at least to Estoup's (1917) classical study.

c. The dependence of α on sample size invalidates Horvath's (1963) claim that α "allows a simple comparison of the results obtained in different empirical situations by noting the variation in this 1 parameter [p. 361]." In the same way, the use of the type–token ratio as a measure for continuous prose has been largely abandoned (cf. Rubenstein & Aborn, 1960), because its decrease with sample size makes comparisons of different-sized samples impossible.

The decreasing α curves for continuous prose data (Figure 2) incidentally raise some questions as to the validity of applying the stochastic model to data from text as well. Simon and Van Wormer (1963) certainly seem to have abandoned the constant-α assumption in favor of the modified assumption of α as either a slowly monotonically decreasing function of k or a step function with different values of α for different portions of the distribution.

The similarity of the α curves for both text and free association data suggests that the same basic mechanisms might be responsible for the generation of words in free association as in continuous discourse. This is consistent with Howes (1957) finding that frequency distributions of free association responses to a number of stimuli are stochastically equivalent to those from large samples of continuous prose.

Comparison of α and R_1

Horvath has claimed that α is a better measure of the associative power of a word than is the frequency of the primary response, R_1 (the associative measure most commonly used). As the following sections show, α and R_1 actually measure different (if not complementary) aspects of free association distributions, and the appropriateness of either depends entirely on which of these aspects is being measured.

What is being measured by α and by R_1? The main advantage of α, as Horvath has correctly pointed out, is that it describes the shape of a free association distribution as a whole. In other words, it is a measure of the breadth of the associative network induced by a particular stimulus. This is something on which R_1, as a measure, does not provide any information.

The value of R_1, on the other hand (as well as of R_2–R_n), is to provide a numerical estimate of the associative strength linking specific pairs of words. Thus R_1 expresses the probability with which a specific word (the most frequent one in a given distribution) can be expected to occur in response to a given stimulus. The empirical usefulness of R_1 has been demonstrated in studies too numerous to mention (e.g., Deese, 1959; Diethorn & Voss, 1967; Jarrett & Scheibe, 1962).

As Horvath has noted, two stimulus words may have the same value of R_1, yet have quite different response distributions. The converse, however, is also true: two response distributions may give the same α value, yet the strengths of the primary responses may be quite different. Thus, in the Montreal norms the stimulus words TABLE and OCÈAN both have α values of .263. Yet R_1 accounts for 57.7% of all responses to TABLE, and only for 14.0% of all responses to OCÈAN.

On the whole, there is a tendency for a reciprocal relationship between R_1 and α, especially when dealing with extreme values of either. Thus, with very large values of R_1, limits will necessarily be set on the maximum possible value of α. For example, since 83.3% of all subjects gave the R_1 "chair" to the stimulus TABLE in the Minnesota norms, only 16.7% of the total responses is left as a potential source of *different* responses. Therefore, the higher the value of R_1, the less likely a high value of α, and the higher the value of α, the less likely a high value of R_1.

Stability with sample size. Horvath has pointed out that the frequency of R_1 is bound to vary from sample to sample. This, however, is only true of the raw frequency of R_1. When raw frequencies are converted into percentages, R_1 can be seen to exhibit quite remarkable constancy over varying sample size. Thus, for the eight samples drawn from the Montreal norms, the mean percentages of R_1 for samples increasing from 25 to 300 were: 19.6, 19.2, 19.2, 19.1, 19.1, 19.2, 19.4, and 19.5, respectively. The differences between these values are so negligible that it becomes clearly superfluous to test for their statistical significance. Thus, contrary to what is implied in Horvath's paper, R_1 is stable across sample size, while α is not.

Stability of the ordinal relationships within a sample. Both R_1 and α show a good deal of variation within any given sample. Some stimuli produce very high R_1 values (e.g., HOMME → 57.7% R_1; LAMPE → 49.0% R_1 in the Montreal norms), while others have very low ones (DOUX → 4.3% R_1; COLÈRE → 6.7% R_1, etc.). The same is true of α, which varies from a low of .206 for the stimulus TABLE to a high of .510 for the stimulus COLÈRE in the Montreal norms. Do these values of α and R_1 reflect characteristics which will remain stable across samples for the same stimulus word?

To test for within-sample stability, the α values for the 100 Kent-Rosanoff stimuli were ranked for each of the eight samples described earlier. Kendall's tau was then calculated for overlapping pairs of samples. All the obtained correlations were highly significant, indicating that the relative values of α remain essentially stable with increasing sample size. (Tau values ranged from .679 for the correlation between samples of $N = 25$ and $N = 50$ subjects, to .914 for the correlation between samples of $N = 250$ and $N = 300$.)

To test for the stability of R_1, a matrix of Pearson correlation coefficients was obtained for the percentage of R_1 values of the same samples. Since R_1 is usually interpreted to mean a dominant association, only stimuli producing the same R_1 in at least five out of the eight samples were included in these calculations. Eighty-five out of the 100 Kent-Rosanoff stimuli were found to satisfy these criteria. The remaining 15 words could be considered, for all practical purposes, as having no primary response at all. They were marked by R_1's very low in numerical value, accompanied by several more or less equiprobable low-frequency R_2's and R_3's.

The cross-correlation matrix of R_1 for the eight samples contained correlations which were highly significant in every instance, with many approaching a perfect linear relationship ($r = .97, .98, .99$). This means that for stimuli having a discernibly dominant R_1, the relative value of this measure can be considered as an extremely stable characteristic of the stimulus word.

Summary and Conclusions

The points raised in this paper may be summarized as follows:

1. Contrary to Horvath's (1963) suggestion, Simon's (1955) stochastic model is not applicable to individual free association hierarchies.

2. The formulae presented by Horvath for calculating expected frequencies are not theoretically justifiable, even if the fit to empirical data may be reasonably good for very large samples (such as the Minnesota norms). Furthermore, even if the model were applicable, it could only provide expected frequencies for the least frequently occurring responses.

3. Because of the inappropriateness of the stochastic model, α is in fact nothing other than the long-familiar type–token ratio, which is known to decrease with increasing sample size.

4. Alpha cannot be directly compared to R_1, as Horvath has done, since each measures different aspects of free association response hierarchies.

5. Because of its dependency on sample size, α is not a good measure for comparing results obtained under different conditions. R_1, on the other hand (if expressed as the percentage of total responses), is extremely invariant with sample size.

6. Both α and R_1 express stable characteristics of stimulus words, since the ordinal relationships of α and R_1 values in any given sample remain very closely the same in samples of different sizes.

REFERENCES

BOUSFIELD, W. A., COHEN, B. H., WHITMARSH, G. A., & KINCAID, W. D. *The Connecticut free association norms.* (Tech. Rep. No. 35, Office of Naval Research Contract Nonr-631(00)) Storrs: University of Connecticut, 1961.

CHOTLOS, J. W. Studies in language behavior: IV. A statistical and comparative analysis of individual written language samples. *Psychological Monographs,* 1944, **56**(2, Whole No. 255).

DEESE, J. A. Influence of inter-item associative strength upon immediate free recall. *Psychological Reports,* 1959, **5**, 305–312.

DIETHORN, J. A., & VOSS, J. F. Serial learning as a function of locus of chained associations. *Journal of Experimental Psychology,* 1967, **73**, 411–418.

ESTOUP, J. B. *Gammes sténographiques.* (5th ed.) Paris, 1917.

FREIBERGS, V. *Normes d'association libre aux 100 mots Kent-Rosanoff.* (Rep. No. 1, Octroi APA-219, Conseil National de Recherches) Montréal: Université de Montréal, 1968.

HORVATH, W. J. A stochastic model for word association tests. *Psychological Review,* 1963, **70**, 361–364.

HOWES, D. On the relation between the probability of a word as an association and in general linguistic usage. *Journal of Abnormal and Social Psychology*, 1957, **54**, 75–85.

JARRETT, R. F., & SCHEIBE, K. E. Association chains and paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1962, **1**, 264–268.

KUČERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English.* Providence: Brown University Press, 1967.

RUBENSTEIN, H., & ABORN, M. Psycholinguistics. *Annual Review of Psychology*, 1960, **11**, 291–322.

RUSSELL, W. A., & JENKINS, J. J. *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test.* (Tech. Rep. No. 11, Contract N8onr-66216, Office of Naval Research and University of Minnesota) Minneapolis: University of Minnesota, 1954.

SIMON, H. A. On a class of skew distribution functions. *Biometrika*, 1955, **42**, 425–440.

SIMON, H. A., & VAN WORMER, T. A. Some Monte Carlo estimates of the Yule distribution. *Behavioral Science*, 1963, **8**, 203–210.

SKINNER, B. F. The distribution of associated words. *Psychological Records*, 1937, **1**, 71–76.

(Received September 9, 1971)

Psychological Review
1972, Vol. 79, No. 3, 275–277

ON THE DISTINCTION BETWEEN DENSITY AND CROWDING:

SOME IMPLICATIONS FOR FUTURE RESEARCH[1]

DANIEL STOKOLS[2]

University of North Carolina

A distinction is drawn between the physical condition, density, defined purely in terms of spatial parameters, and the experience of crowding, a motivational state aroused through the interaction of spatial, social, and personal factors, and directed toward the alleviation of perceived spatial restriction. The implications of this distinction for future research are discussed.

To date, there have been four basic lines of behavioral research which relate to the issue of crowding: (*a*) animal studies (Calhoun, 1962, 1966; Christian, Flyger, & Davis, 1960), (*b*) correlational surveys utilizing census tract data (Chombart de Lauwe, 1959; Mitchell, 1971; Schmitt, 1957, 1966; Winsborough, 1965), (*c*) experiments on the human use of space (Barker, 1965, 1968; Hall, 1959, 1966; Sommer, 1967, 1969), and (*d*) experimental studies directly concerned with the effects of crowding on human behavior (Freedman, 1970; Freedman, Klevansky, & Ehrlich, 1971; Griffit & Veitch, 1971; Hutt & Vaizey, 1966; Ittelson, Proshansky, & Rivlin, 1970).

Although previous empirical approaches provide some insights into the nature of crowding phenomena, there appears to be a certain confusion, reflected especially in research concerned with human crowding, regarding the meaning of the terms "density" and "crowding." Many writers have used these terms interchangeably, rather than distinguishing between the physical condition, density, involving spatial limitation, and the experiential state, crowding, in which the restrictive aspects of limited space are perceived by the individuals exposed to them. As a consequence of this confusion, there has been a general tendency to view crowding in terms of spatial considerations alone, and a failure to delineate those social and personal dimensions which may interact with spatial factors to mediate the experience of crowding.

According to the proposed distinction, density is viewed as a necessary antecedent, rather than a sufficient condition, for the experience of crowding. Any instance of spatial limitation involves potential inconveniences—the restriction of movement or the preclusion of privacy, for example. These potential constraints, however, are not necessarily salient to the individuals occupying an area of limited space. While the amount of space in a given area may appear limited to an outside observer, it will not inevitably seem inadequate to the occupants of the area, especially if their activities do not require a high degree of behavioral coordination, if their relationships with each other are cooperative and friendly, or if they have had much experience with living and working under conditions of limited space. Such circumstances, then, would operate to minimize the salience of spatial constraints.

In order for the restrictive aspects of spatial limitation to become salient and aversive, such that the individual's perception of them engenders the experience of crowding, certain social and personal factors must be present.[3] Social interference, for example, in the form of task-coordination problems or competition with others, is likely to heighten the salience of spatial constraints. The restriction of movement, imposed by limited space, would become most apparent while engaging in tasks requiring the coordination of one's own activity with that of others; and the arousal of competitive feelings would eventuate in a tendency to perceive the presence of others as threatening and intrusive, thereby generating a desire to expand and protect one's personal space. Similarly, the lack of prior experience with spatial limitation, as well as certain characterological traits (e.g., impatience, aggressiveness), would render an individual more susceptible to the experience of crowding.

---

[1] The preparation of this paper was supported by United States Public Health Service Grant 5-TO1 MH 07325. The author wishes to express his appreciation to John Schopler and John Thibaut for their critical readings of the manuscript. Thanks are also due to Nehemia Friedland for his comments on this paper.

[2] Requests for reprints should be sent to Daniel Stokols, Department of Psychology, University of North Carolina, Chapel Hill, North Carolina 27514.

[3] To the extent that spatial limitation introduces noxious physical effects, such as heightened temperature and stuffiness, the relative importance of social and personal determinants of crowding decreases. Hence, under conditions of acute spatial limitation, density alone would be sufficient to induce the experience of crowding. The present discussion, however, focuses on situations in which the supply of space is not limited to the point that it impairs normal biological functioning, or promotes excessive physical discomfort.

Crowding, then, appears to arise through the juxtaposition of density with certain social and personal circumstances which sensitize the individual to the potential constraints of limited space. The perception of such constraints leads to a recognized disparity between the amount of space demanded, or considered to be adequate, by the individual, and the amount of space available to him. The experience of crowding, thus, can be characterized as a motivational state directed toward the alleviation of perceived restriction and infringement, through the augmentation of one's supply of space, or the adjustment of social and personal variables so as to minimize the inconveniences imposed by spatial limitation.

The particular form of one's response to crowding will be a function of the relative intensity of spatial, social, and personal factors, and the degree to which they can be modified. Where the limitation of space is extreme, and restraints against direct alteration of spatial variables are low, the prepotent mode of response to crowding will be a behavioral one. For instance, the individual can augment his supply of personal space by leaving the crowded situation. In situations where either normative or physical constraints inhibit overt behavioral adjustments of spatial variables, perceptual and cognitive modes of reducing the salience of restricted space will be more likely to occur. In such cases, the person may modify his standards of spatial adequacy, enhance the attractiveness of his task, or attempt to achieve a greater degree of coordination with others in the group, as a means of alleviating the sensation of crowding.

Although unsuccessful attempts to cope with crowding may be accompanied by behaviors symptomatic of general stress (e.g., discomfort, aggression, hormonal disequilibrium), crowding is distinguishable, nonetheless, from other stress syndromes by the fact that it involves a prevailing concern with spatial constraints and the motivation to eliminate them, or reduce their salience.

The suggested distinction between density and crowding has important implications for future research. Recent experiments have attempted to manipulate crowding through variations in the size of the experimental room (Freedman, 1970; Freedman et al., 1971). Moreover, these studies have reported very few experimental effects on the dependent assessments of task performance and affective behavior. Though interpretation of negative results is tenuous, at best, the conceptualization of crowding, discussed above, suggests that the lack of experimental effects in previous investigations may be due to their failure to manipulate, or control for, personal and social sources of variance. Though the variation of experimental room size may well represent an effective manipulation of available space per subject, it does not appear to constitute an adequate manipulation of the determinants of crowding. Through the orthogonal manipulation of variables pertaining to social and personal dimensions, however, as well as those relating to the spatial component, it becomes possible to assess variance attributable to the former factors which, heretofore, has been subsumed under "error variance."

While previous investigations have focused on density as an independent variable, subsequent research must turn its attention to the experience of crowding as a measurable phenomenon. Such an approach would permit direct assessment of the spatial, social, and personal parameters of crowding, through the simultaneous manipulation of them. Moreover, the behavioral, perceptual, and cognitive consequences of perceived crowding could be assessed through the use of (*a*) behavioral measures pertaining to the augmentation of one's space, (*b*) subjective reports of restriction and discomfort, (*c*) observational indexes of tension (e.g., laughter, aggression, reduced eye contact), and (*d*) physiological indicators of strain. The utilization of such assessments would contribute to an elucidation of the crowding experience, as distinct from other motivational syndromes.

REFERENCES

Barker, R. G. Explorations in ecological psychology. *American Psychologist*, 1965, **20**, 1–14.

Barker, R. G. *Ecological psychology*. Stanford: Stanford University Press, 1968.

Calhoun, J. B. Population density and social pathology. *Scientific American*, 1962, **206**, 139–148.

Calhoun, J. B. The role of space in animal sociology. *Journal of Social Issues*, 1966, **22**, 46–59.

Chombart de Lauwe, P. *Famille et habitation*. Paris: Editions du Centre National de la Recherche Scientific, 1959.

CHRISTIAN, J., FLYGER, V., & DAVIS, D. Factors in the mass mortality of a herd of sika deer *Cervus nippon*. *Chesapeake Science*, 1960, 1, 79–95.

FREEDMAN, J. The effects of crowding on human behavior. Unpublished manuscript, Department of Psychology, Columbia University, 1970.

FREEDMAN, J., KLEVANSKY, S., & EHRLICH, P. The effect of crowding on human task performance. *Journal of Applied Social Psychology*, 1971, **1**, 7–25.

GRIFFIT, W., & VEITCH, R. Hot and crowded: Influences of population density and temperature on interpersonal affective behavior. *Journal of Personality and Social Psychology*, 1971, **17**, 92–98.

HALL, E. *The silent language.* Greenwich, Conn.: Premier Books, 1959.

HALL, E. *The hidden dimension.* Garden City, N. Y.: Doubleday, 1966.

HUTT, C., & VAIZEY, M. Differential effects of group density on social behavior. *Nature*, 1966, **209**, 1371–1372.

ITTELSON, W., PROSHANSKY, H., & RIVLIN, L. The environmental psychology of the psychiatric ward. In H. Proshansky, W. Ittelson, & L. Rivlin (Eds.), *Environmental psychology.* New York: Holt, Rinehart & Winston, 1970.

MITCHELL, R. Some implications of high density housing. *American Sociological Review*, 1971, **36**, 18–29.

SCHMITT, R. C. Density, delinquency and crime in Honolulu. *Sociology and Social Research*, 1957, **41**, 274–276.

SCHMITT, R. C. Density, health, and social disorganization. *Journal of the American Institute of Planners*, 1966, **32**, 38–40.

SOMMER, R. Small group ecology. *Psychological Bulletin*, 1967, **67**, 145–152.

SOMMER, R. *Personal space—The behavioral basis of design.* Englewood Cliffs, N. J.: Prentice Hall, 1969.

WINSBOROUGH, H. The social consequences of high population density. *Law and Contemporary Problems*, 1965, **30**, 120–126.

(Received October 21, 1971)

Psychological Review
1972, Vol. 79, No. 3, 278–279

A COMMENT ON D. J. MOORE AND D. A. SHIEK'S "TOWARD A THEORY OF EARLY INFANTILE AUTISM"[1]

ROGER A. WEBB[2]

Johns Hopkins University

D. J. Moore and D. A. Shiek's theory concerning the origins of infantile autism is examined and found implausible. The theory is seen as based on the misuse or misunderstanding of two key concepts—imprinting and intelligence. Their notion of imprinting is mechanical and does not take into account the necessary environmental interaction which is characteristic of social development in man and other mammals. More serious is Moore and Shiek's belief that intellectual potential would be indicative of developmental precocity. Extensive empirical study indicates that there is no relation between the behavior of infants at birth and later IQ. It is suggested that all theories of autism based on the presumed precocity of children be abandoned as unproductive.

In their recent paper on early infantile autism, Moore and Shiek (1971) hypothesize that: "fetuses with an inherited high potential for intellectual superiority are developmentally ready for imprinting with human behaviors much before birth . . . [and become] imprinted on the restricted uterine environment [p. 451]." Their theory deserves comment as a potential source of confusion. While novel theories are clearly needed in the difficult and important field of early infantile autism, it is unlikely that Moore and Shiek's formulation will prove useful. Their paper implies a notion of development which does not do justice to the complexities of the process. While other points might be examined, this note will concentrate on Moore and Shiek's misunderstanding or misapplication of two fundamental concepts—imprinting and intelligence.

The realization that early human social development is based on instinctive processes that are related to similar mechanisms in other animals is an insight of profound importance (cf. Bowlby, 1969, for a review). However, as Bowlby and other have shown, primary socialization in humans (as well as other mammals) occurs well after birth and depends on extensive environmental interaction. For example, considerable perceptual organization must occur before the infant can begin to distinguish strange from familiar faces. Also, several relatively independent systems must be organized and focused on important social figures. Considered in these terms, the socialization process is better compared to Piaget's analysis of the construction of the permanent object than to elicited following in ducks. Also, Moore and Shiek fail to note that social responsiveness persists with no sharp cutoff for some time after the six-week to six-month period they postulate (Bowlby, 1969, p. 321 ff.). Thus, it is unlikely that any aspect of the normal human socialization process could function in the uterine environment to produce the profound and lasting effects of autism.

Moore and Shiek's misconceptions about the nature and development of intelligence are even more serious. They write: "If his [the autistic child's] eventual IQ is projected to be 170 or more his unstimulated brain has already [in utero] passed through the first six months of what would ordinarily be his postnatal life [p. 455]." It is a plausible hypothesis (though not well established) that autistic children have (or had) considerable intellectual potential, but the projection of an IQ of 170+ is totally unfounded. Moreover, their logic would seem to imply that all children with potential IQs of 170+ should be autistic, and that all children with $100 < IQ < 170$ should be ready for "imprinting" before the period of normal socialization. Such great variability in early development is, in fact, not seen, and there are nonautistic, highly intelligent people about.

These considerations indicate that Moore and Shiek hold an outmoded and incorrect notion of what intelligence is. First, there is an implicit assumption that social precocity could be predicted from a measure of IQ or mental age, and this is not likely to be true. Second, there is an erroneous assumption that presumed intellectual potential would have some implications for the state of the infant at the time of birth or before.

---

[1] Preparation of this note was supported in part by Grant MH-19364 from the National Institute of Mental Health. The author wishes to thank Natalie Hirsch and Julian Stanley for discussing the issues, though neither takes responsibility for the ideas expressed.

[2] Requests for reprints should be sent to Roger A. Webb, Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218.

The idea that an IQ taken at any age can give an accurate mental age at any other age holds only if mental age/chronological age is the proper conceptualization of intelligence and if IQ is constant at all ages. Neither prerequisite is likely to be correct. One could argue that there is no necessary relationship between precocity and intelligence and that, in fact, the two notions are conceptually unrelated. There is, of course, an empirical relation between the acquisition of intellectual material and intelligence in the years of childhood. Regardless of the truth of this argument, however, one can safely say that there is no demonstrated relationship between behaviors measured in the first year or so of life and later intelligence, despite a long and extensive search for such relations (Bayley, 1970, p. 1173 ff.). In fact, factor-analytic studies suggest that the abilities which eventually contribute the majority of variance to intelligence do not emerge clearly until age four or so (Hofstaetter, 1954). Thus, even if autistic children were of high intellectual potential, there would be no grounds for assuming there was anything functionally different about them at birth.

Realization of these limitations on theorizing is important, since Moore and Shiek are not alone in confusing precocity and later intelligence. Rimland (1964), for example, bases his early brain-damage hypothesis on an assumption of intellectual potential. He postulates that the more rapidly developing brain of the highly intelligent child is more likely to suffer injury at birth. Moore and Shiek's rejection of this theory on an argument of biological adaptability is incorrect, since the brain is more susceptible to damage when it is undergoing rapid change. Rimland's hypothesis is gratuitous, however, for the same reasons that applied to Moore and Shiek's above. It is to be hoped that all theories of autism that attribute a critical role to the infant's intellectual potential will be abandoned, and that research on this difficult problem will examine potentially more profitable areas.

REFERENCES

BAYLEY, N. Development of mental abilities. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology*. (3rd ed.) New York: Wiley, 1970.

BOWLBY, J. *Attachment and loss*. Vol. 1. *Attachment*. New York: Basic Books, 1969.

HOFSTAETTER, P. R. The changing composition of intelligence: A study in T-technique. *Journal of Genetic Psychology*, 1954, **85**, 159–164.

MOORE, D. J., & SHIEK, D. A. Toward a theory of infantile autism. *Psychological Review*, 1971, **78**, 451–456.

RIMLAND, B. *Infantile autism*. New York: Appleton-Century-Crofts, 1964.

(Received December 16, 1971)

Psychological Review
1972, Vol. 79, No. 3, 280

A RESPONSE TO R. A. WEBB'S COMMENTS

DEWEY J. MOORE[1]

Indiana State University

Webb (1972) accepts as a plausible hypothesis that autistic children have (or had) high intellectual potential. He also agrees that normal human social development has its basis in instinctive processes and is preceded by considerable perceptual organization. He further agrees that such perceptual organization is unlikely to occur in utero. Having agreed with all the major tenets of the theory, Webb proceeds to argue its conclusions by citing truisms from normal development as his evidence. He conveniently ignores the fact that early infantile autism is not a normal state.

Webb anticipates that if the theory is valid there should be degrees of autism. He notes correctly that such variability is not observed. The IQ limit of 170+ was not chosen capriciously or arbitrarily. As hypothesized in the original paper, a projected intellectual level of such magnitude is necessary for the critical six months to pass in utero. Less potential permits adequate opportunity for imprinting and/or perceptual organization after birth. There are probably time limits for imprinting in humans as in lower animals. After expiration of these limits, the developmental opportunity is lost. Imprinting seems to be an all-or-nothing-at-all phenomenon. In autism, it is hypothesized to be a nothing-at-all phenomenon. Without such basic perceptual organization, the human socialization process fails to proceed normally. Hence, Webb is correct in his observation that early infantile autism does not exist in degrees, but he is correct for reasons other than those he states.

Finally, Webb rejects the theory's rejection of Rimland's (1964) early brain damage hypothesis on grounds that the brain is more susceptible to damage when it is developing rapidly. His restatement of this commonly accepted truism lends nothing to the discussion. Significantly he failed to comment on the fact that the negative effects of sensory deprivation are relatively greater for genetically brighter organisms.

The theory demands flexibility of thought, rigidity of logic, and a renunciation of many time-worn and erroneous beliefs regarding autism. Hopefully, experimenters and theoreticians will scrutinize past, present, and future knowledge regarding the status of autistics in the light of new hypotheses. One avenue of exploration might attempt to determine how biochemical differences that appear to exist in the reticular activating systems of autistics approximate those found in normals who have undergone sensory dprivation.

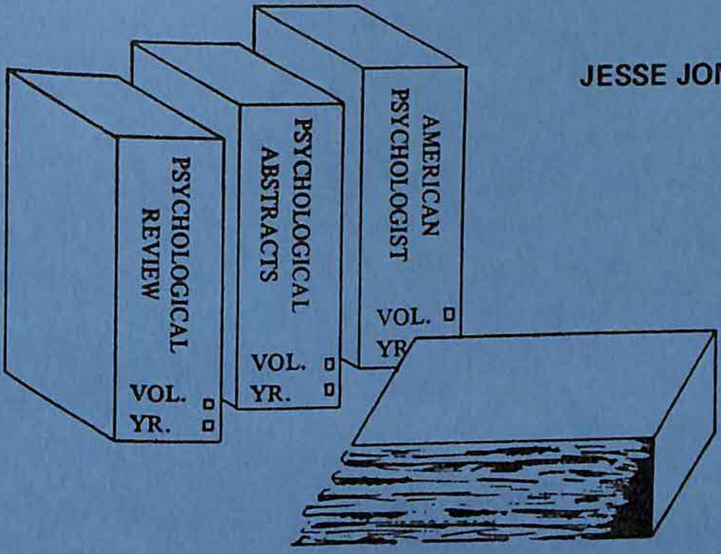
REFERENCES

Rimland, B. *Infantile autism.* New York: Appleton-Century-Crofts, 1964.

Webb, R. A. A comment on D. J. Moore and D. A. Skiek's "Toward a theory of early infantile autism." *Psychological Review,* 1972, **79,** 278–279.

---

[1] Requests for reprints should be sent to Dewey J. Moore, Department of Special Education, Indana State University, Terre Haute, Indiana 47809.



JESSE JONES VOLUME FILES

Send orders to:
JESSE JONES CORP.
P.O. Box 5120
Philadelphia, Pa. 19141

Only $3.50 ea.
3 for $10.00
6 for $19.00 P.P.D.

Now, the journals of the American Psychological Association can become an attractive permanent part of your professional library. These famous Jesse Jones volume files, especially designed to keep your copies orderly, readily accessible for future references – guard against soiling, tearing, wear or misplacement of copies.

This durable file will support 150 lbs. Looks and feels like leather and is washable. The 23-carat gold lettering makes it a fit companion for the most costly binding.

2 FEB 1974

The British Journal of Psychology

Volume 63 Part I February 1972

CONTENTS

LESLIE HENDERSON On mental energy

HANS KREITLER and SHULAMITH KREITLER The model of cognitive orientation: towards a theory of human behaviour

JAMES G. TAYLOR Underwater distortion: and plain distortion

JOHN COHEN, E. I. CHESNICK and D. HARAN A confirmation of the inertial-ψ effect in sequential choice and decision

JOHN N. NICHOLSON and JEFFREY A. GRAY Peak shift, behavioural contrast and stimulus generalization as related to personality and development in children

JOHN M. INNES Word-association response commonality and the generation of associative structures

VERNON HAMILTON The size constancy problem in schizophrenia: a cognitive skill analysis

DIANA R. HASLAM Field dependence in relation to pain threshold

MARGARET A. VINCE and SUSAN CHINN Effects of external stimulation on the domestic chick's capacity to stand and walk

RODERICK WONG Infantile handling and associative processes of rats

KIM KIRSNER Developmental changes in short-term recognition memory

PAPERS TO APPEAR IN FORTHCOMING ISSUES

BOOK REVIEWS

£2.50 net ($8.50 in USA). Annual subscription £7.50 net ($25.00 in USA) for four parts

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 EUSTON ROAD, LONDON NW1 2DB
American Branch: 32 East 57th Street, New York, NY 10022

KC
to Dockaris
R the Library
1/2/74

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 79, No. 4 July 1972

Articles

Elimination by Aspects: A Theory of Choice
Amos Tversky 281

Absolute Judgment and Paired-Associate Learning: Kissing Cousins or Identical Twins?
Jane A. Siegel and William Siegel 300

Selective Attention and the Breadth of Learning: An Extension of the One-Look Model
Bryan E. Shepp, Deborah G. Kemler, and Daniel R. Anderson 317

A Chronometric Analysis of Simple Addition
Guy J. Groen and John M. Parkman 329

Laws of the Simple Visual Reaction Time
Warren H. Teichner and Marjorie J. Krebs 344

Subjective Contours and Apparent Depth
Stanley Coren 359

Letters to the Editor 368

Editorial Note 316

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

CHRIS ARGYRIS
A. H. BLACK
DONALD S. BLOUGH
ROBERT C. BOLLES
LYLE E. BOURNE
GORDON H. BOWER
JACK W. BREHM
PETER E. BRYANT
MICHAEL COLE
SEYMOUR FESHBACH
JOHN M. FOLEY
JOHN GIBBON
JAMES J. GIBSON
J. A. GRAY
RICHARD L. GREGORY
ROBERT G. GRICE
CHARLES S. HARRIS
JULIAN HOCHBERG
EARL B. HUNT
LEO M. HURVICH
JAMES JOHNSTON
LAWRENCE KARLIN
SEYMOUR LEVINE
ELIJAH LOVEJOY
R. DUNCAN LUCE
N. J. MACKINTOSH
DOUGLAS MEDIN
JACOB NACHMIAS
RAYMOND S. NICKERSON
ALBERT PEPITONE
IRWIN POLLACK
DAVID RAAB
FRANK RESTLE
PAUL ROZIN
MARTIN E. P. SELIGMAN
LARRY STEIN
P. SUPPES
A. TREISMAN
VICTOR H. VROOM

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts submitted to the *Psychological Review*. In each issue the Editor will list in the *Review* those editorial consultants who have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowledged.)

The *Psychological Review* is devoted to articles of theoretical significance to any area of scientific endeavor in psychology. Except for occasional articles, manuscripts exceeding 7,500 words will be discouraged. Ordinarily, manuscripts which consist primarily of original reports of research or of reviews of the literature should be submitted to other journals. Methodological and statistical articles are not appropriate for the *Psychological Review*. See the editorials in the *Review* (1965, **72**, 1–2, 1971, **78**, 1–2) for further details.

Beginning with the 1971 issue, Letters to the Editor will be published on an experimental basis. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obtained from the APA Central Office.

All manuscripts must be submitted in triplicate. In the reference list, give journal titles in full; do not abbreviate. Figures should be submitted as glossy prints; duplicate figures may be photographic or pencil-drawn copies. Authors should retain a copy of the manuscript in case there is loss in the mail.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Manuscripts should be addressed to the Editor, George Mandler, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of materials copyrighted by the American Psychological Association: Written permission must be obtained for copying or reprinting any tables, figures, or text of more than 500 words in length. Persons or organizations who obtain permission to reproduce any such copyrighted materials must include on the first page of the reproduced materials the exact copyright notice which appeared on the Association's original publication. There will be a charge of $10.00 per page for such permission except when the requester is the author of the article to be reprinted, or when the article is to be reproduced in a limited number of copies solely for instructional purposes.

HAROLD P. VAN COTT
Managing Editor

LORRAINE BIROS
Technical
Editor

HENRY S. SALMON
Advertising and Business
Manager

ANNE REDMAN
Subscription Manager

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1200 Seventeenth St., N. W., Washington, D. C. 20036

$10.00 volume

Second-class postage paid at Lancaster, Pa.

$2.00 issue

Send all communications, including address changes, to 1200 Seventeenth St., N. W., Washington, D. C. 20036. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PSYCHOLOGICAL
REVIEW

Diary No 727(L)
Date 4.2
Researc

Vol. 79, No. 4 JULY 1972

ELIMINATION BY ASPECTS: A THEORY OF CHOICE[1]

AMOS TVERSKY[2]

Hebrew University of Jerusalem

Most probabilistic analyses of choice are based on the assumption of simple scalability which is an ordinal formulation of the principle of independence from irrelevant alternatives. This assumption, however, is shown to be inadequate on both theoretical and experimental grounds. To resolve this problem, a more general theory of choice based on a covert elimination process is developed. In this theory, each alternative is viewed as a set of aspects. At each stage in the process, an aspect is selected (with probability proportional to its weight), and all the alternatives that do not include the selected aspect are eliminated. The process continues until all alternatives but one are eliminated. It is shown (*a*) that this model is expressible purely in terms of the choice alternatives without any reference to specific aspects, (*b*) that it can be tested using observable choice probabilities, and (*c*) that it generalizes the choice models of R. D. Luce and of F. Restle. Empirical support from a study of psychophysical and preferential judgments is presented. The strategic implications of the present development are sketched, and the logic of elimination by aspects is discussed from both psychological and decision-theoretical viewpoints.

When faced with a choice among several alternatives, people often experience uncertainty and exhibit inconsistency. That is, people are often not sure which alternative they should select, nor do they always make the same choice under seemingly identical conditions. In order to account for the observed inconsistency and the reported uncertainty, choice behavior has been viewed as a probabilistic process.

Probabilistic theories of preference differ with respect to the nature of the mechanism that is assumed to govern choice. Some theories (e.g., Thurstone, 1927, 1959) attribute a random element to the determination of subjective value, while others (e.g., Luce, 1959) attribute a random element to the decision rule. Most theoretical work on probabilistic preferences has been based on the notion of independence among alternatives. This notion, however, is incompatible with some observed patterns of preferences which exhibit systematic dependencies among alternatives.

This paper develops a probabilistic theory of choice, based on a covert elimination process, which accounts for observed

---

[1] The research was supported, in part, by National Science Foundation Grant GB-6782. Much of the work reported in this paper was accomplished while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, California, during 1970–1971. I wish to thank the Center for the generous hospitality. I am grateful to David H. Krantz for many invaluable discussions throughout the years, to Maya Bar-Hillel for her assistance in both theoretical and experimental phases of the investigation, and to Edward N. Pugh for his help in the analysis of the data. I have also benefited from discussions with Clyde H. Coombs, Robyn M. Dawes, R. Duncan Luce, Jacob Marschak, J. E. Russo, and Paul Slovic.

[2] Requests for reprints should be sent to Amos Tversky, who is now at the Oregon Research Institute, P. O. Box 3196, Eugene, Oregon 97403.

© 1972 by the American Psychological Association, Inc.

dependencies among alternatives. The first section analyzes the independence assumption; the second section formulates a theory of choice and discusses its consequences; some experimental tests of the theory are reported in the third section; and its psychological implications are explored in the fourth and final section.

We begin by introducing some notation. Let $T = \{x,y,z,\cdots\}$ be a finite set, interpreted as the total set of alternatives under consideration. We use A, B, C, $\cdots$, to denote specific nonempty subsets of T, and A_i, B_j, C_k, $\cdots$, to denote variables ranging over nonempty subsets of T. Thus, $\{A_i | A_i \supseteq B\}$ is the set of all subsets of T which includes B. The number of elements in A is denoted by a. Proper and nonproper set inclusion are denoted, respectively, by $\supset$ and $\supseteq$. The empty set is denoted by ϕ. The probability of choosing an alternative x from an offered set $A \subseteq T$ is denoted $P(x,A)$. Naturally, we assume $P(x,A) \geq 0$, $\sum_{x \in A} P(x,A) = 1$ for any A, and $P(x,A) = 0$ for any $x \notin A$. For brevity, we write $P(x;y)$ for $P(x,\{x,y\})$, $P(x;y,z)$ for $P(x,\{x,y,z\})$, etc. A real-valued, nonnegative function in one argument is called a *scale*. Choice probability is typically estimated by relative frequency in repeated choices. It should be kept in mind, however, that other empirical interpretations of choice probability, such as confidence judgments (which are applicable to unique choice situations), might also be adopted.

Perhaps the most general formulation of the notion of independence from irrelevant alternatives is the assumption that the alternatives can be scaled so that each choice probability is expressible as a monotone function of the scale values of the respective alternatives. This assumption, called *simple scalability*, was first investigated by Krantz (1964, Appendix A). Formally, simple scalability holds if and only if there exists a scale u defined on the alternatives of T and functions F_n in n arguments, $2 \leq n \leq t$, such that for any $A = \{x,\cdots,z\} \subseteq T$,

$$P(x,A) = F_a[u(x),\cdots,u(z)], \qquad [1]$$

where each F_a is strictly increasing in the first argument and strictly decreasing in the remaining $a - 1$ arguments provided $P(x,A) \neq 0, 1$. This assumption underlies most theoretical work in the field. The theory of Luce (1959), for example, is a special case of this assumption where

$$P(x,A) = F_a[u(x),\cdots,u(z)] = \frac{u(x)}{\sum_{y \in A} u(y)}. \qquad [2]$$

Despite its generality, simple scalability (Equation 1) has strong testable consequences. In particular, it implies that for all x, $y \in A$,

$$P(x;y) \geq 1/2 \text{ iff } P(x,A) \geq P(y,A), \quad \text{provided } P(y,A) \neq 0. \qquad [3]$$

Equation 3 asserts that the ordering of x and y, by choice probability, is independent of the offered set.[3] Thus, if x is preferred to y in one context (e.g., $P(x;y) \geq 1/2$), then x is preferred to y in any context. Furthermore, if $P(x;y) = 1/2$ then $P(x,A) = P(y,A)$ for any A which contains both x and y. Thus, if an individual is indifferent between x and y, then he should choose them with equal probability from any set which contains them.

This assumption, however, is not valid in general, as suggested by several counterexamples and demonstrated in many experiments (see Becker, DeGroot, & Marschak, 1963b; Chipman, 1960; Coombs, 1958; Krantz, 1967; Tversky & Russo, 1969). To motivate the present development, let us examine the arguments against simple scalability starting with an example proposed by Debreu (1960).

Suppose you are offered a choice among the following three records: a suite by Debussy, denoted D, and two different recordings of the same Beethoven symphony, denoted B_1 and B_2. Assume that the two Beethoven recordings are of equal quality, and that you are undecided between adding a Debussy or a Beethoven to your record collection. Hence, $P(B_1;B_2) = P(D;B_1) = P(D;B_2) = 1/2$. It follows readily from Equation 3 that $P(D;B_1,B_2) = 1/3$. This conclusion, however, is unacceptable on intuitive grounds because the basic conflict between Debussy and Beethoven is not likely to be affected by the addition of another Beethoven recording. Instead, it is suggested that in choosing among the three records, B_1 and B_2 are treated as one alternative to be compared with D. Consequently, one would expect that $P(D;B_1,B_2)$ will be close to one-half, while $P(B_1;B_2,D) = P(B_2;B_1,D)$ will be close to one-fourth, contrary to simple scalability (Equation 1). Empirical support for Debreu's hypothesis was presented by Becker et al. (1963b) in a study of choice among gambles. Although Debreu's example was offered as a criticism of Luce's model (Equation 2), it applies to any model based on simple scalability.

---

[3] Simple scalability is, in fact, equivalent (see Tversky, 1972) to the following *order independence* assumption. For x, $y \in A - B$, and $z \in B$, $P(x,A) \geq P(y,A)$ iff $P(z,B \cup \{x\}) \leq P(z,B \cup \{y\})$ provided the terms on the two sides of either inequality are not both 0 or 1.

Previous efforts to resolve this problem (e.g., Estes, 1960) attempted to redefine the alternatives so that B_1 and B_2 are no longer viewed as different alternatives. Although this idea has some appeal, it does not provide a satisfactory account of our problem. First, B_1 and B_2 are not only physically distinct, but they can also be perfectly discriminable. Hence, there is no independent basis for treating them as indistinguishable. Second, the process of redefining choice alternatives itself requires an adequate theoretical analysis. Finally, data show that the principle of independence from irrelevant alternatives is violated in a manner that cannot be readily accounted for by grouping choice alternatives. More specifically, it appears that the addition of an alternative to an offered set "hurts" alternatives that are similar to the added alternative more than those that are dissimilar to it. Such an effect (of which Debreu's example is a special case) requires a more drastic revision of the principles underlying our models of choice.

The following example provides another illustration of the inadequacy of simple scalability. Suppose each of two travel agencies, denoted 1 and 2, offers tours of Europe (E) and of the Far East (F). Let $T = \{E_1,F_1,E_2,F_2\}$ where letters denote the destination of the tours, and the subscripts denote the respective agencies. Let us assume, for simplicity, that the decision maker is equally attracted by Europe and by the Far East, and that he has no reason to prefer one travel agency over the other. Consequently, all binary choice probabilities equal one-half, and the probability of choosing each tour from the total set equals one-fourth. It follows from Equation 3, in this case, that all trinary probabilities must equal one-third. However, an examination of the problem suggests that in fact none of the trinary probabilities equals one-third; instead, some of them equal one-half while the others equal one-fourth.

Consider, for example, the set $\{E_1,F_1,F_2\}$. Since the distinction between the agencies is treated as irrelevant, the problem reduces to the choice between a tour of Europe and a tour of the Far East. If the latter is chosen, then either one of the agencies can be selected. Consequently, $P(E_1;F_1,F_2) = 1/2$, and $P(F_1;F_2,E_1) = P(F_2;F_1,E_1) = 1/4$. An identical argument applies to all other triples. Besides violating simple scalability, this example demonstrates that the same set of binary (or quarternary) probabilities can give rise to different trinary probabilities and hence the latter cannot be determined by the former. Put differently, this example shows that the probabilities of choosing alternatives from a given set, A, cannot be computed, in general, from the probabilities of choosing these alternatives from the subsets and the supersets of A. This observation imposes a high lower bound on the complexity of any adequate theory of choice.

A minor modification of an example due to L. J. Savage (see Luce & Suppes, 1965, pp. 334–335), which is based on binary comparisons only, illustrates yet another difficulty encountered by simple scalability. Imagine an individual who has to choose between a trip to Paris and a trip to Rome. Suppose he is indifferent between the two trips so that $P(\text{Paris}; \text{Rome}) = 1/2$. When the individual is offered a new alternative which consists of the trip to Paris plus a \$1 bonus, denoted Paris +, he will undoubtedly prefer it over the original trip to Paris with certainty so that $P(\text{Paris}+; \text{Paris}) = 1$. It follows from Equation 3, then, that $P(\text{Paris}+; \text{Rome}) = 1$, which is counterintuitive. For if our individual cannot decide between Paris and Rome, it is unlikely that a relatively small bonus would resolve the conflict completely and change the choice probability from 1/2 to 1. Rather, we expect $P(\text{Paris}+; \text{Rome})$ to be closer to 1/2 than to 1. Experimental data (e.g., Tversky & Russo, 1969) support this intuition. Choice probabilities, therefore, reflect not only the utilities of the alternatives in question, but also the difficulty of comparing them. Thus, an extreme choice probability (i.e., close to 0 or 1) can result from either a large discrepancy in value or from an easy comparison, as in the case of the added bonus. The comparability of the alternatives, however, cannot be captured by their scale values, and hence simple scalability must be rejected. The above examples demonstrate that the substitution of one alternative for another, which is equivalent to it in some contexts, does not necessarily preserve choice probability in any context. The substitution affects the comparability among the alternatives, which in turn influences choice probability.

An alternative approach to the development of probabilistic theories of choice treats the utility of each alternative as a random variable rather than a constant. Specifically, it is assumed that there exists a random vector $\mathbf{U} = (\mathbf{U}_x, \cdots, \mathbf{U}_z)$ on $T = \{x, \cdots, z\}$ (i.e., for any $y \in T$, $\mathbf{U}_y$ is a random variable) such that

$$P(x,A) = P(\mathbf{U}_x \geq \mathbf{U}_y \text{ for all } y \in A). \quad [4]$$

Models of this type are called random utility models. The only random utility models which have been seriously investigated assume that the random variables are independent. However, an extension of the last example (see Luce & Raiffa, 1957, p. 375) is shown to violate any independent random utility model. To demonstrate, consider the trips to Paris and Rome with and without the added bonus. The expected binary choice probabilities in this case are $P(\text{Paris}+; \text{Paris}) = 1$, $P(\text{Rome}+; \text{Rome}) = 1$ but $P(\text{Paris}+; \text{Rome}) < 1$ and $P(\text{Rome}+; \text{Paris}) < 1$.

Assuming an independent random utility model, the first two equations above imply that there is no overlap between the distributions representing Paris and Paris +, nor is there an overlap between the distributions representing Rome and Rome +. The last two inequalities above imply that there must be some overlap between the distributions representing Rome and Paris +, as well as between the distributions representing Paris and Rome +. It is easy to verify that these conclusions are mutually inconsistent, and hence the above choice probabilities are incompatible with any independent random utility model. The representation of choice alternatives by independent random variables, therefore, appears too restrictive in general since, like simple scalability, it is incompatible with some eminently reasonable patterns of preference. In discussing the difficulties encountered by probabilistic theories of choice, Luce and Suppes (1965) wrote:

> It appears that such criticisms, although usually directed toward specific models, are really much more sweeping objections to all our current preference theories. They suggest that we cannot hope to be completely successful in dealing with preferences until we include some mathematical structure over the set of outcomes that are simply substitutable for one another and those that are special cases of others. Such functional and logical relations among the outcomes seem to have a sharp control over the preference probabilities, and they cannot long be ignored [p. 337].

THEORY

The present development describes choice as a covert sequential elimination process. Suppose that each alternative consists of a set of aspects of characteristics,[4] and that at every stage of the process, an aspect is selected (from those included in the available alternatives) with probability that is proportional to its weight. The selection of an aspect eliminates all the alternatives that do not include the selected aspect, and the process continues until a single alternative remains. If a selected aspect is included in all the available alternatives, no alternative is eliminated and a new aspect is selected. Consequently, aspects that are common to all the alternatives under consideration do not affect choice probabilities. Since the present theory describes choice as an elimination process governed by successive selection of aspects, it is called the elimination-by-aspects (EBA) model.

---

[4] The representation of choice alternatives as collections of measurable aspects was developed by Restle (1961) who formulated a binary choice model based on this representation. As will be shown later, the present theory reduces to Restle's in the

In contemplating the purchase of a new car, for example, the first aspect selected may be automatic transmission: this will eliminate all cars that do not have this feature. Given the remaining alternatives, another aspect, say a \$3000 price limit, is selected and all cars whose price exceeds this limit are excluded. The process continues until all cars but one are eliminated. This decision rule is closely related to the lexicographic model (see Coombs, 1964; Fishburn, 1968), where an ordering of the relevant attributes is specified a priori. One chooses, then, the alternative that is best relative to the first attribute; if some alternatives are equivalent with respect to the first attribute, one chooses from them the alternative that is best relative to the second attribute, and so on. The present model differs from the lexicographic model in that here no fixed prior ordering of aspects (or attributes) is assumed, and the choice process is inherently probabilistic.

More formally, consider a mapping that associates with each $x \in T$ a nonempty set $x' = \{\alpha,\beta,\cdots\}$ of elements which are interpreted as the aspects of x. An alternative x is said to include an aspect α whenever $\alpha \in x'$. The aspects could represent values along some fixed quantitative or qualitative dimensions (e.g., price, quality, comfort), or they could be arbitrary features of the alternatives that do not fit into any simple dimensional structure. The characterization of alternatives in terms of aspects is not necessarily unique. Furthermore, we generally do not know what aspects are considered by an individual in any particular choice problem. Nevertheless, as is demonstrated later, this knowledge is not required in order to apply the present model, and its descriptive validity can be determined independently of any particular characterization of the alternatives.

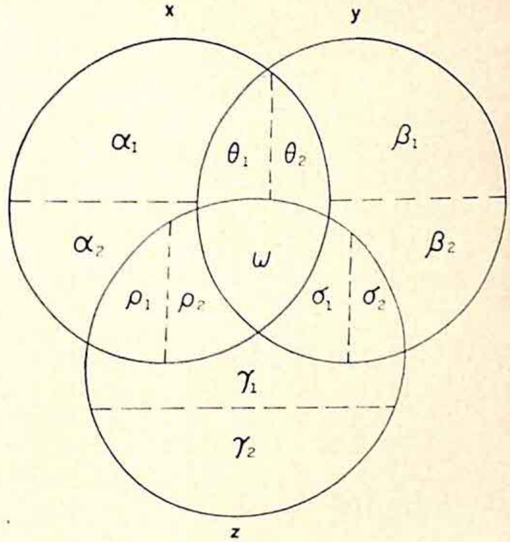


FIG. 1. A graphical representation of aspects in the three-alternative case.

To clarify the formalization of the model, let us first examine a simple example. Consider a three-alternative set $T = \{x,y,z\}$, where the collections of aspects associated with the respective alternatives are

$$x' = \{\alpha_1,\alpha_2,\theta_1,\theta_2,\rho_1,\rho_2,\omega\},$$

$$y' = \{\beta_1,\beta_2,\theta_1,\theta_2,\sigma_1,\sigma_2,\omega\},$$

and

$$z' = \{\gamma_1,\gamma_2,\rho_1,\rho_2,\sigma_1,\sigma_2,\omega\}.$$

A graphical representation of the structure of the alternatives and their aspects is presented in Figure 1. It is readily seen that α_i, β_i, and γ_i $(i = 1, 2)$ are, respectively, the unique aspects of x, y, and z; that θ_i, σ_i, and ρ_i are, respectively, the

---

two-alternative case. A related representation of choice alternatives was developed by Lancaster (1966) who assumed that economic goods possess, or give rise to, multiple characteristics (or aspects) in fixed proportion, and that these characteristics determine the consumer's choice. Lancaster's theory, however, is nonprobabilistic.

aspects shared by x and y, by y and z, and by x and z; and that ω is shared by all three alternatives. Since the selection of ω does not eliminate any alternative, it can be discarded from further considerations. Let u be a scale which assigns to each aspect a positive number representing its utility or value, and let K be the sum of the scale values of all the aspects under consideration, that is, $\mathrm{K} = \sum_{\alpha} u(\alpha)$ where the summation ranges over all the aspects except ω. Using these notations we now compute $P(x,T)$.

Note first that x can be chosen directly from T if either α_1 or α_2 is selected in the first stage (in which case both y and z are eliminated). This occurs with probability $[u(\alpha_1) + u(\alpha_2)]/\mathrm{K}$. Alternatively, x can be chosen via $\{x,y\}$ if either θ_1 or θ_2 is selected in the first stage (in which case z is eliminated), and then x is chosen over y. This occurs with probability $[u(\theta_1) + u(\theta_2)] \times P(x;y)/\mathrm{K}$. Finally, x can be chosen via $\{x,z\}$ if either ρ_1 or ρ_2 is selected in the first stage (in which case y is eliminated), and then x is chosen over z. This occurs with probability $[u(\rho_1) + u(\rho_2)]P(x;z)/\mathrm{K}$. Since the above paths leading to the choice of x from T are all disjoint,

$$\begin{aligned} P(x,T) = (1/\mathrm{K})(&u(\alpha_1) + u(\alpha_2) \\ &+ [u(\theta_1) + u(\theta_2)]P(x;y) \\ &+ [u(\rho_1) + u(\rho_2)]P(x;z)) \end{aligned} \qquad [5]$$

where

$$P(x;y) = \frac{u(\alpha_1) + u(\alpha_2) + u(\rho_1) + u(\rho_2)}{u(\alpha_1) + u(\alpha_2) + u(\rho_1) + u(\rho_2) + u(\beta_1) + u(\beta_2) + u(\sigma_1) + u(\sigma_2)}, \text{ etc.}$$

More generally, let T be any finite set of alternatives. For any $A \subseteq T$ let $A' = \{\alpha \mid \alpha \in x' \text{ for some } x \in A\}$, and $A^0 = \{\alpha \mid \alpha \in x' \text{ for all } x \in A\}$. Thus, A' is the set of aspects that belongs to at least one alternative in A, and A^0 is the set of aspects that belongs to all the alternatives in A. In particular, T' is the set of all aspects under consideration, while T^0 is the set of aspects shared by all the alternatives under study. Given any aspect $\alpha \in T'$, let A_α denote those alternatives of A which include α, that is, $A_\alpha = \{x \mid x \in A \ \& \ \alpha \in x'\}$.

The elimination-by-aspects model asserts that there exists a positive scale u defined on the aspects (or more specifically on $T' - T^0$) such that for all $x \in A \subseteq T$

$$P(x,A) = \frac{\sum_{\alpha \in x' - A^0} u(\alpha)P(x,A_\alpha)}{\sum_{\beta \in A' - A^0} u(\beta)} \qquad [6]$$

provided the denominator does not vanish. Note that the summations in the numerator and the denominator of Equation 6 range, respectively, over all aspects of x and A except those that are shared by all elements of A. Hence, the denominator of Equation 6 vanishes only if all elements of A share the same aspects, in which case it is assumed that $P(x,A) = 1/a$.

Equation 6 is a recursive formula. It expresses the probability of choosing x from A as a weighted sum of the probabilities of choosing x from the various subsets of A (i.e., A_α for $\alpha \in x'$), where the weights (i.e., $u(\alpha)/\sum u(\beta)$) correspond to the probabilities of selecting the respective aspects of x.

Consider a special case of the elimination-by-aspects model where all pairs of alternatives share the same aspects, that is, $x' \cap y' = z' \cap w'$ for all x, y, z, $w \in T$. Since aspects that are common to all the alternatives of T do not affect the choice process, the alternatives can be treated as (pairwise) disjoint, that is, $x' \cap y' = \phi$ for all x, $y \in T$. In this case, Equation 6 reduces to

$$P(x,A) = \frac{\sum_{\alpha \in x'} u(\alpha)}{\sum_{\beta \in A'} u(\beta)}$$

since $\alpha \in x'$ implies $A_\alpha = \{x\}$, and $P(x,\{x\}) = 1$. Letting

$$u(x) = \sum_{\alpha \in x'} u(\alpha)$$

yields

$$P(x,A) = \frac{u(x)}{\sum_{y \in A} u(y)}.$$

Hence, in the present theory, Luce's model

(Equation 2) holds whenever the alternatives can be regarded as composed of disjoint aspects.

Next, examine another special case of the model where only binary choice probabilities are considered. In this case, we obtain

$$P(x;y) = \frac{\sum_{\alpha \in x' - y'} u(\alpha)}{\sum_{\alpha \in x' - y'} u(\alpha) + \sum_{\beta \in y' - x'} u(\beta)}$$

$$= \frac{u(x' - y')}{u(x' - y') + u(y' - x')}, \quad [7]$$

where $x' - y' = \{\alpha \mid \alpha \in x' \ \&\ \alpha \notin y'\}$ is the set of aspects that belongs to x but not to y; $y' - x' = \{\beta \mid \beta \in y' \ \&\ \beta \notin x'\}$ is the set of aspects that belongs to y but not to x; and $u(x' - y') = \sum_{\alpha \in x' - y'} u(\alpha)$. Equation 7 coincides with Restle's (1961) model. The EBA model, therefore, generalizes the choice models of Luce and of Restle.

The elimination-by-aspects model has been formulated above in terms of a scale u defined over the set of relevant aspects. It appears that the application of the model presupposes prior characterization of the alternatives in terms of their aspects. However, it turns out that this is not necessary because the EBA model can be formulated purely in terms of the alternatives, or more specifically, in terms of the subsets of T.

To illustrate the basic idea, consider the example presented in Figure 1. There we assume that α_i, β_i, $\cdots$ $(i = 1, 2)$ are all distinct aspects. According to the elimination-by-aspects model, however, there is no need to distinguish between aspects that lead to the same outcome. For example, the selection of either α_1 or α_2 eliminates both y and z; the selection of either θ_1 or θ_2 eliminates z; and the selection of either ρ_1 or ρ_2 eliminates y. From the standpoint of the elimination-by-aspects model, therefore, there is no need to differentiate between α_1 and α_2, between θ_1 and θ_2, or between ρ_1 and ρ_2. Thus we can group all the aspects that belong to x alone, all the aspects that belong to x and y but not to z, etc. Let $\{\bar{x}\}$ denote the aspects that belong to x alone (i.e., α_1 and α_2), $\{\overline{x,y}\}$ the aspects that belong only to x and y (i.e., θ_1 and θ_2), $\{\overline{x,z}\}$ the aspects that belong only to x and z (i.e., ρ_1 and ρ_2), etc.[5] The representation of the grouped aspects in the three-alternative case is displayed in Figure 2.

The scale value of a collection of aspects is defined as the sum of the scale value of its members, that is, $U(\bar{x}) = u(\alpha_1) + u(\alpha_2)$, $U(\overline{x,y}) = u(\theta_1) + u(\theta_2)$, etc. For simplicity of notation we write $U(\bar{x})$ for $U(\{\overline{x}\})$, $U(\overline{x,y})$ for $U(\{\overline{x,y}\})$, etc. Thus, Equation 5 is expressible as

$$P(x,T) = \frac{U(\bar{x}) + U(\overline{x,y})P(x;y) + U(\overline{x,z})P(x;z)}{U(\bar{x}) + U(\bar{y}) + U(\bar{z}) + U(\overline{x,y}) + U(\overline{x,z}) + U(\overline{y,z})} \quad [8]$$

where

$$P(x;y) = \frac{U(\bar{x}) + U(\overline{x,z})}{U(\bar{x}) + U(\bar{y}) + U(\overline{x,z}) + U(\overline{y,z})},$$

etc.

The essential difference between Equations 5 and 8 lies in the domain of the scales: in Equation 5, u is defined over individual aspects, whereas in Equation 8 U is defined over collections of aspects which are associated, respectively, with the subsets of T. The method by which Equation 5 is translated into Equation 8 can be applied in general.

Each proper subset A of T is associated with the set $\bar{A}$ of all aspects that are included in *all* the alternatives of A and are not included in *any* of the alternatives that do not belong to A. That is, $\bar{A} = \{\alpha \in T' \mid \alpha \in x'$ for all $x \in A$ & $\alpha \notin y'$ for any $y \notin A\}$. The scale U is defined by $U(\bar{A}) = \sum_{\alpha \in \bar{A}} u(\alpha)$. It is shown in the appendix that the elimination-by-aspects model, defined in Equation 6, holds if and only if there exists a scale U defined on $\{\bar{A}_i \mid A_i \subset T\}$ such

---

[5] In this paper, the superbar is used exclusively to denote collections of aspects. It should not be confused with a common use of this symbol to denote set complement.

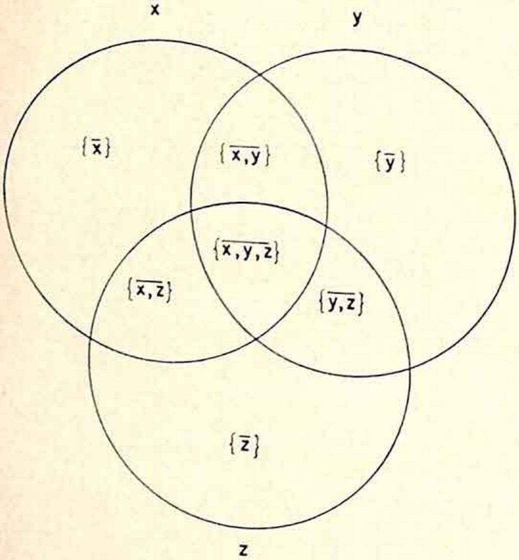


FIG. 2. A graphical representation of the grouped aspects in the three-alternative case.

that for all $x \in A \subseteq T$

$$P(x,A) = \frac{\sum_{B_i \nsupseteq A} U(\bar{B}_i)P(x, A \cap B_i)}{\sum_{A_j \in \mathfrak{A}} U(\bar{A}_j)} \qquad [9]$$

where $\mathfrak{A} = \{A_j | A_j \cap A \neq A, \phi\}$, provided the denominator does not vanish. (According to the present theory, the denominator can vanish only if $P(x,A) = 1/a$.) The significance of this result lies in showing how the elimination-by-aspects model can be formulated in terms of the subsets of T without reference to specific aspects. Note that for $A \subset T$, $U(\bar{A})$ is *not* a measure of the value of the alternatives of A; rather it is a measure of all the evaluative aspects that are shared by all the alternatives of A and by them only. Thus, $U(\bar{A})$ can be viewed as a measure of the unique advantage of the alternatives of A. The reader is invited to verify that in the three-alternative case, Equation 9 reduces to Equation 8.

Before discussing the consequences of the EBA model, let us examine how it resolves the counterexamples described in the previous section. First, consider Debreu's record selection problem where $T = \{D, B_1, B_2\}$. Naturally, the two Beethoven recordings have much more in common with each other than either of them has with the Debussy record. Assume, for simplicity, that any aspect shared by D and one of the B records is also shared by the other B record, hence D can be treated as (aspectwise) disjoint of both B_1 and B_2. Suppose $U(\bar{B}_1) = U(\bar{B}_2) = a$, $U(\overline{B_1, B_2}) = b$, and $U(\bar{D}) = a + b$. A graphical illustration of this representation is shown in Figure 3.

It follows readily, under these assumptions, that all the binary choice probabilities are equal, since

$$P(B_1; B_2) = \frac{a}{2a} = \frac{1}{2} = \frac{a+b}{2(a+b)} = P(D; B_1) = P(D; B_2),$$

yet the trinary choice probabilities are unequal, since

$$P(D; B_1, B_2) = \frac{a+b}{3a+2b} > \frac{a + b(a/2a)}{3a+2b} = P(B_1; B_2, D) = P(B_2; B_1, D).$$

In fact, as a (or a/b) approaches 0, the left-hand side approaches 1/2 while the right-hand side approaches 1/4. Hence, according to the elimination-by-aspects model, all three records can be pairwise equivalent, and yet the probability of choosing D from the entire set can be as high as 1/2 whenever B_1 and B_2 include the same aspects.

Second, consider Savage's problem of choosing between trips, and let $T = \{P, R, P+, R+\}$, where P and R denote, respectively, trips to Paris and Rome, while + denotes a small monetary bonus. Here it is natural to suppose that Paris + includes Paris (in the sense that all aspects of the

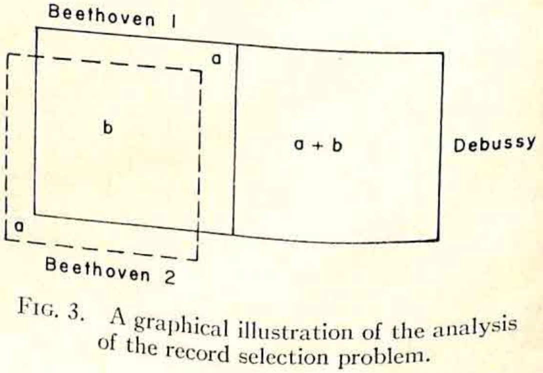


FIG. 3. A graphical illustration of the analysis of the record selection problem.

latter trip are included in the former). On the other hand, Paris + does not include Rome because each of these trips has some aspects that are not shared by the other. Similarly, Rome + includes Rome but not Paris. The relations among the four alternatives are illustrated in Figure 4.

Letting $U(\overline{\mathrm{P}+}) = U(\overline{\mathrm{R}+}) = a$, and $U(\overline{\mathrm{P},\mathrm{P}+}) = U(\overline{\mathrm{R},\mathrm{R}+}) = b$, yields

$$P(\mathrm{P};\mathrm{R}) = \frac{b}{2b} = \frac{1}{2} = \frac{a+b}{2(a+b)} = P(\mathrm{P}+;\mathrm{R}+),$$

$$P(\mathrm{P}+;\mathrm{P}) = P(\mathrm{R}+;\mathrm{R}) = \frac{a}{a} = 1, \quad \text{and}$$

$$P(\mathrm{P}+;\mathrm{R}) = P(\mathrm{R}+;\mathrm{P}) = \frac{a+b}{a+2b},$$

which can take any value between 1/2 and 1, depending on the relative weight of the bonus. Thus, the above pattern of binary choice probabilities, which violates simple scalability (Equation 1) and any independent random utility model (Equation 4), arises naturally in the present model. Essentially the same solution to this problem (which involves only binary probabilities) has been proposed by Restle (1961).

The reader is invited to show how the elimination-by-aspects model can accommodate the example described earlier of choice among tours of Europe or the Far East with each of two travel agencies.

Consequences

In the following discussion we assume that the elimination-by-aspects model is valid, and list some of its testable consequences. The derivations of these properties are presented in Tversky (1972).

Regularity: For all $x \in A \subseteq B$,

$$P(x,A) \geq P(x,B). \quad [10]$$

Regularity asserts that the probability of choosing an alternative from a given set cannot be increased by enlarging the offered set. This is probably the weakest form of noninteraction among alternatives. Although regularity seems innocuous, it is worth noting that the replacement of $\geq$ by $>$ in Equation 10 violates the expected preference pattern in the record selection problem.

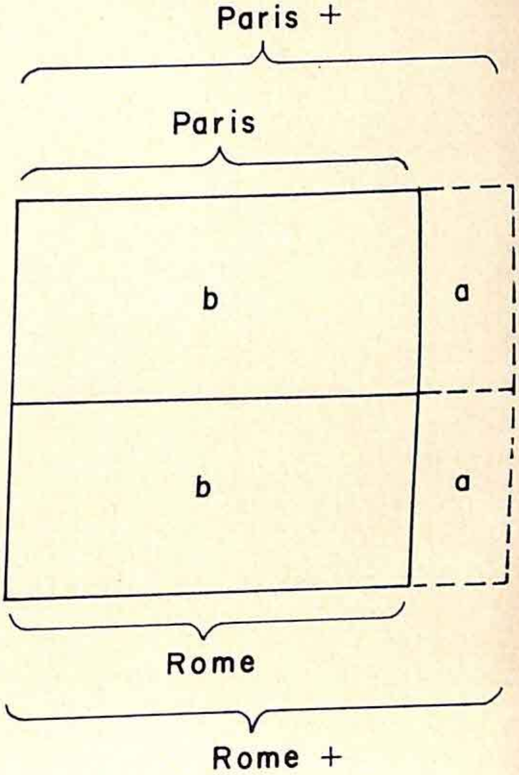


FIG. 4. A graphical illustration of the analysis of the choice between trips.

The following consequence of the elimination-by-aspects model involves binary probabilities only. Since it generalizes the algebraic notion of transitivity, it is called *moderate stochastic transitivity*.

Moderate stochastic transitivity:

$$P(x;y) \geq 1/2 \text{ and } P(y;z) \geq 1/2 \text{ imply}$$
$$P(x;z) \geq \min[P(x;y), P(y;z)]. \quad [11]$$

If we replace min by max in the conclusion of Equation 11, we obtain a stronger condition called *strong stochastic transitivity*. This latter property (which is not a consequence of the present model) is essentially equivalent to simple scalability in the binary case. If we replace the conclusion of Equation 11 by $P(x;z) \geq 1/2$, we obtain a weaker condition called *weak stochastic transitivity*, which is a consequence of the existence of an ordinal utility scale satisfying $u(x) \geq u(y)$ iff $P(x;y) \geq 1/2$.

The next consequence of the EBA model has not been investigated previously to the best of my knowledge. It relates binary and trinary choice probabilities by a property called the *multiplicative inequality.*

Multiplicative inequality:

$$P(x;y,z) \geq P(x;y)P(x;z). \quad [12]$$

The multiplicative inequality asserts that the probability of choosing x from $\{x,y,z\}$ is at least as large as the probability of choosing x from both $\{x,y\}$ and $\{x,z\}$ in two independent choices. It is conjectured that the elimination-by-aspects model implies a much stronger form of the multiplicative inequality, namely, $P(x,A \cup B) \geq P(x,A)P(x,B)$ for all A, $B \subseteq T$.

Equations 10 and 12 can be combined to yield

$$\min[P(x;y), P(x;z)] \geq P(x;y,z) \geq P(x;y)P(x;z). \quad [13]$$

Thus, trinary choice probabilities are bounded from above by regularity, and from below by the multiplicative inequality. A geometric representation of Equation 13 which displays the admissible range of $P(x;y,z)$ given the values of $P(x;y)$ and $P(x;z)$ is given in Figure 5. It shows that the trinary probability must lie between the lower and upper surfaces generated, respectively, by the multiplicative inequality (Equation 12) and regularity (Equation 10).

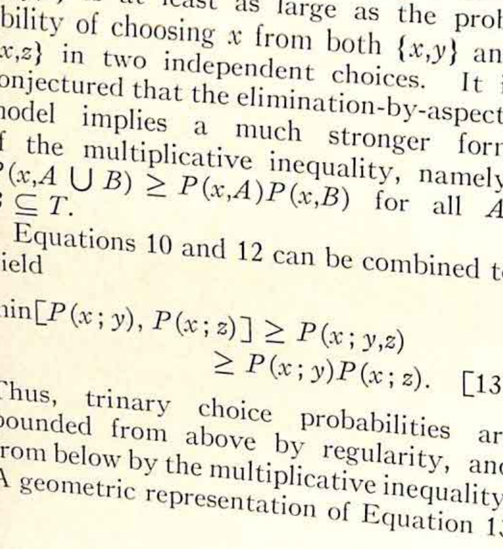


FIG. 5. A geometric representation of the admissible values (shaded region) of the trinary probability $P(x;y,z)$ given the binary probabilities $P(x;y)$ and $P(x;z)$, under Equation 13.

The significance of the above consequences stems from the fact that they provide measurement-free tests of the elimination-by-aspects model, that is, tests which do not require estimation of parameters.

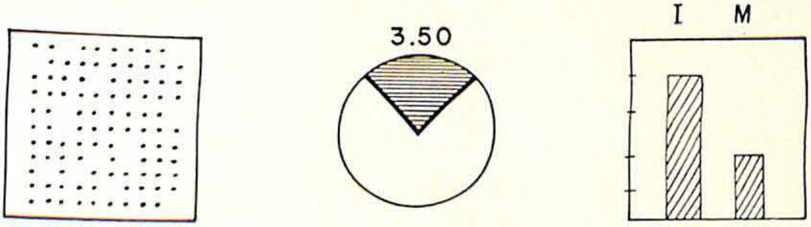
For a given set of alternatives T, the elimination-by-aspects model has $2^t - 3$ free parameters, or U values (the number of proper nonempty subsets of T minus an arbitrary unit of measurement), while the number of independent data points of the form $P(x,A)$, $x \in A \subseteq T$, is

$$\sum_{n=2}^{t} (n-1)\binom{t}{n} = (t-2)2^{t-1} + 1.$$

Hence, there are always at least as many data points as parameters in the present model; the former exceeds the latter whenever $t > 3$. In general, therefore, the scale values are uniquely determined by the choice probabilities except in some particular situations, for example, when $P(x,A) = 1/a$ for all $x \in A \subseteq T$.

Even in the case where $t = 3$, in which the number of parameters (five) equals the number of data points, the choice probabilities are severely constrained. The volume of the subspace generated by the present model is less than $1/2\%$ of the volume of the entire parameter space which is a five-dimensional unit hypercube. The probability that a point sampled at random, from a uniform distribution over the parameter space, satisfies the present model, therefore, is less than .005 in this case.

Additional consequences and further developments of the elimination-by-aspects model are presented in Tversky (1972). They include a generalization of the present model, an extension to ranking, and a proof that the EBA model is a random utility model, though not an independent one.



A dot pattern **A gamble** **A score profile**

FIG. 6. Typical stimulus slides from each of the three tasks.

TESTS

In contrast to the many theoretical studies of probabilistic models of preference (see, e.g., Becker et al., 1963a; Luce & Suppes, 1965; Marschak, 1960; Morrison, 1963), there have been relatively few empirical studies in which these models were tested. Moreover, much of the available data are limited to binary choices, and most studies report and analyze only group data (see, e.g., Rumelhart & Greeno, 1971). Unfortunately, group data usually do not permit adequate testing of theories of individual choice behavior because, in general, the compatibility of such data with the theory is neither a necessary nor a sufficient condition for its validity. (For an instructive illustration of this point, see Luce, 1959, p. 8.) The scarcity of appropriate data in an area of considerable theoretical interest is undoubtedly due to the difficulties involved in obtaining adequate estimates of choice probabilities for an individual subject, particularly outside the domain of psychophysics.

Two consequences of the present model were tested in previous studies. In an experiment involving choice among gambles, Becker et al. (1963b) showed that although simple scalability (Equation 1) is systematically violated, the regularity condition (Equation 10) is generally satisfied. Similarly, although strong stochastic transitivity was violated in several studies (e.g., Coombs, 1958; Krantz, 1967; Tversky & Russo, 1969), moderate stochastic transitivity was usually supported. (For some specified conditions under which moderate stochastic transitivity, as well as weak stochastic transitivity, is violated, see Tversky, 1969.) The fact that simple scalability and strong stochastic transitivity are often violated while regularity and moderate stochastic transitivity are typically satisfied provides some support, albeit nonspecific, for the present theory. The following experimental work was designed to obtain a more direct test of the EBA model.

Method

To test the model, three different tasks were selected. The stimuli in Task *A* were random dot patterns, in a square frame, varying in size (of square) and density (of dots). Subjects were presented with pairs and triples of frames and instructed to choose, in each case, the frame which contained the largest number of dots. The stimuli in Task *B* were profiles of college applicants with different intelligence (I) and motivation (M) scores. The scores were expressed in percentiles (relative to the population of college applicants), and displayed as bar graphs. Subjects were presented with pairs and triples of such profiles and asked to select, in each case, the applicant they considered the most promising. The stimuli in Task *C* were two-outcome gambles of the form (p,x), in which one wins $\$x$ with probability p and nothing otherwise. Each gamble was displayed as a pie diagram, where the probabilities of winning and not winning were represented, respectively, by the black and white sectors of the pie. Subjects were presented with pairs and triples of gambles and were asked to choose the gamble they would prefer to play. (At the end of the study, each subject actually played for money five of the gambles chosen by him in the course of the study. The gambles were played by spinning an arrow on a wheel of fortune and the subjects won the indicated amont if the arrow landed on the black sector of the wheel.) Examples of the three types of stimuli are shown in Figure 6.

The same eight subjects participated in all three tasks. They were students in a Jerusalem high school, ages 16–18. Subjects were run in a single group. The stimuli were projected on slides and each subject indicated his choices by checking an appropriate box on his response sheet. The study consisted of 12 one-hour sessions, three times a week, for four weeks. The first two sessions were practice sessions in which the problems and the procedure were introduced and the subjects familiarized themselves with the stimuli of the task.

Each experimental session included all three tasks, and the ordering of the tasks was randomized across sessions. Within each task, subjects were presented with various pairs and triples formed from a basic set of 4 × 4 = 16 two-dimensional stimuli. One set of three stimuli of each type was isolated and replicated more than other sets. The entire triple was replicated 30 times (three per session) while each of the pairs within this triple was replicated 20 times (two per session). The following discussion is concerned with the analysis of these triples. Each triple was constructed so that no alternative dominates another one with respect to both dimensions, and so that two of the elements, called x and y, are very similar to each other, while the third element, z, is relatively dissimilar to each of them.[6]

The subjects were paid a flat fee for the completion of all the sessions. In addition, each subject received a bonus proportional to the number of correct numerosity judgments made by him, and was allowed to play, for money, five gambles selected randomly from those chosen by him during the study.

Results

The analysis of the results begins by testing the constant-ratio rule which is essentially equivalent to Luce's (1959) model. According to this rule,

$$\frac{P(x;y)}{P(y;x)} = \frac{P(x,A)}{P(y,A)} \quad x, y \in A, \qquad [14]$$

provided the denominators do not vanish. The constant-ratio rule is a strong version of the principle of independence from irrelevant alternatives. It requires that the ratio of $P(x,A)$ and $P(y,A)$ (not merely their order as required by simple scalability) be independent of the offered set A.

Let $T = \{x,y,z\}$, and define

$$P_y(x;z) = \frac{P(x,T)}{P(x,T) + P(z,T)},$$

$$P_x(y;z) = \frac{P(y,T)}{P(y,T) + P(z,T)}.$$

Hence, by the constant-ratio rule,

$$P(x;z) = P_y(x;z)$$

and

$$P(y;z) = P_x(y;z). \qquad [15]$$

Put differently, the binary probability $P(x;z)$ should equal $P_y(x;z)$, computed from the trinary probabilities, since under Equation 14 the presence of y is "irrelevant" to the choice between x and z.

In the present study, the alternatives were designed so that x and y are much more similar to each other than either of them is to z. Hence, the similarity hypothesis that is incorporated into the elimination-by-aspects model predicts that the addition of alternative y to the set $\{x,z\}$ will reduce $P(x,T)$ proportionally more than $P(z,T)$. That is, the similar alternative, x, will lose relatively more than the dissimilar alternative, z, by the addition of y. Likewise, y is expected to lose relatively more than z by the introduction of x. Contrary to the constant-ratio rule, therefore, the similarity hypothesis implies

$$P(x;z) > P_y(x;z)$$

and

$$P(y;z) > P_x(y;z). \qquad [16]$$

To test the constant-ratio rule, the observed (binary) relative frequencies $\hat{P}(x,z)$ and $\hat{P}(y,z)$ were compared, respectively, with $\hat{P}_y(x;z)$ and $\hat{P}_x(y;z)$ computed from the trinary relative frequencies, separately for each one of the subjects. The observed and the computed values for all subjects are shown in Table 1 for each of the three tasks.

It seems that the constant-ratio model (Equation 14) holds in the psychophysical task (A), and that it fails in the two preference tasks (B and C) in the manner predicted by the similarity hypothesis (Equation 16). Out of 16 individual comparisons in each task (two per subject), Equation 16 was satisfied in 13 and 15 cases, respec-

---

[6] The following stimuli were employed in the study. Task $A: x = (13 \times 13, 4/5)$, $y = (14 \times 14, 3/4)$, and $z = (28 \times 28, 1/5)$ where the first component of each stimulus is the size of the underlying matrix used to generate the pattern, and the second component is the proportion of cells of the matrix that contain dots. Task $B: x = (78,25)$, $y = (75,35)$, and $z = (60,90)$ where the first and second components of each pair denote, respectively, intelligence and motivation scores of the applicants. Task C: $x = (1/5, 4.00)$, $y = (1/4, 3.50)$, and $z = (2/3, 1.00)$, where the first and second component of each pair are, respectively, the probability of winning and the amount to be won in each of the gambles in Israeli pounds.

TABLE 1

OBSERVED AND PREDICTED PROPORTIONS (UNDER THE CONSTANT-RATIO MODEL) FOR EACH TASK

Subject	Task *A* (dots)				Task *B* (applicants)				Task *C* (gambles)			
	$\hat{P}(x;z)$	$\hat{P}_y(x;z)$	$\hat{P}(y;z)$	$\hat{P}_z(y;z)$	$\hat{P}(x;z)$	$\hat{P}_y(x;z)$	$\hat{P}(y;z)$	$\hat{P}_z(y;z)$	$\hat{P}(x;z)$	$\hat{P}_y(x;z)$	$\hat{P}(y;z)$	$\hat{P}_z(y;z)$
1	.50	.43	.45	.43	.65	.44	.30	.26	.35	.12	.50	.46
2	.60	.27	.35	.33	.55	.37	.75	.58	.60	.53	.70	.68
3	.25	.38	.40	.41	.55	.38	.60	.41	.25	.26	.50	.29
4	.70	.75	.30	.67	.40	.46	.40	.32	.60	.43	.70	.35
5	.65	.52	.35	.39	.65	.45	.55	.40	.20	.16	.50	.41
6	.40	.39	.45	.52	.35	.20	.40	.38	.65	.54	.60	.44
7	.15	.26	.45	.44	.75	.77	.35	.40	.55	.42	.65	.50
8	.15	.14	.45	.57	.55	.52	.40	.29	.55	.35	.70	.43
Overall proportion	.425	.405	.400	.466	.556	.463	.469	.388	.469	.354	.606	.466
p	*ns*		*ns*		$< .05$		$< .10$		$< .01$		$< .01$	

tively, in Tasks *B* and *C* ($p < .05$ in each case[7]), and only in 7 cases in Task *A*. Essentially the same result was found in additional analyses.

The relatively small number of observations does not permit an adequate test of individual comparisons. Hence, the observed and the computed choice frequencies were pooled over subjects. The results of a chi-square test of Equation 15 against Equation 16, based on these data, are shown in the last row of Table 1 for each comparison in each of the tasks. The same pattern emerges from the analysis of the pooled data: the observed proportions are significantly higher than the computed ones in Tasks *B* and *C*, but not in Task *A*.

Since the constant-ratio model is not acceptable, in general, the simplest version of the elimination-by-aspects model, which is compatible with the similarity hypothesis, was selected next. Recall that the test stimuli were designed so that x and y are very similar to each other while z is relatively dissimilar to either of them (see Footnote 6). Thus, we assume that neither x nor y share with z any aspect that they do not share with each other. Consequently, aside from the aspects shared by all three stimuli, z can be regarded as (aspectwise) disjoint from both x and y. That is, we assume that, to a reasonable degree of approximation, $U(\overline{x,z}) = U(\overline{y,z}) = 0$. This assumption reduces the number of free parameters (from five to three) at the cost of some loss in generality.

Let $U(\bar{x}) = a$, $U(\bar{y}) = b$, $U(\bar{z}) = c$, and $U(\overline{x,y}) = d$ (see Figure 7). Under this special case of the model, there exist nonnegative a, b, c, and d such that

$$P(x;y) = \frac{a}{a+b}, \quad P(y;z) = \frac{b+d}{b+d+c},$$

$$P(x;z) = \frac{a+d}{a+d+c}, \tag{17}$$

$$P(x;y,z) = \frac{a + d\dfrac{a}{a+b}}{a+b+c+d},$$

and

$$P(z;x,y) = \frac{c}{a+b+c+d}.$$

For three alternatives, there are five independent data points (three binary and two trinary). In the absence of any restrictions on the parameters, the likelihood function of the data is maximized by using the observed relative frequencies as estimates of the parameters, in which case the dimensionality of the parameter space, denoted $d(\Omega)$, equals five. In the above version (Equation 17) of the elimination-by-aspects model, we can set c, say, arbitrarily, whence the observed proportions are all expressible in terms of three parame-

---

[7] This significance level should be interpreted with caution because of the potential dependency between the observations of each subject.

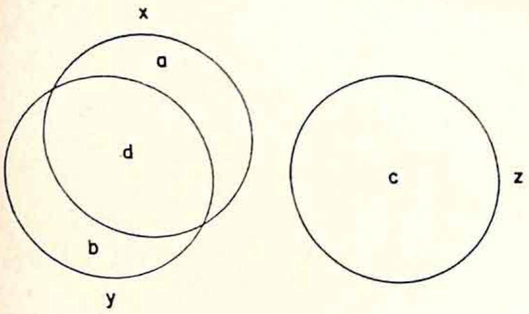


FIG. 7. A graphical illustration of the tested version of the EBA model (Equation 17).

ters (a, b, and d), and the dimensionality of the restricted parameter space, denoted $d(\omega)$, equals three. Let λ be the likelihood ratio $L(\omega)/L(\Omega)$, where L denotes the maximum value of the likelihood function under the respective model. If Equation 17 holds, then the statistic $-2ln\lambda$ has an approximate chi-square distribution with $d(\Omega) - d(\omega) = 2$ degrees of freedom.

Chandler's (1969) STEPIT program was employed to obtain maximum likelihood estimates of the parameters under Equation 17 with $c = 1$. The values of the test statistics are reported in Table 2, along with the estimates of d, for each subject in all tasks.

Table 2 exhibits a very good correspondence between the observed proportions and the tested version (Equation 17) of the EBA model: only 2 out of 24 tests permit rejecting the model at the conservative .1 level. It should perhaps be noted that a correspondence between observed choice probabilities and the elimination-by-aspects model does not necessarily imply that the subjects are actually following a strategy of elimination by aspects. They might, in fact, employ a different strategy that is well approximated by the elimination-by-aspects model. The study of the actual strategies employed by subjects in choice experiments may perhaps be advanced by investigating choice probabilities in conjunction with other data such as reaction time, eye movements, or verbal protocols.

TABLE 2

VALUES OF THE TEST STATISTIC AND THE ESTIMATED VALUES OF d FOR EACH SUBJECT IN EACH OF THE TASKS

Subject	Task A (dots)		Task B (applicants)		Task C (gambles)	
	χ^2	d	χ^2	d	χ^2	d
1	.133	.29	2.179	.14	.040	.46
2	3.025	.89	1.634	.92	.001	.58
3	.849	0	.159	1.18	2.022	.14
4	5.551*	0	6.864*	.51	1.053	1.56
5	.951	0	.428	1.23	.887	0
6	.401	0	.405	.42	.157	1.18
7	3.740	0	.083	0	.304	1.00
8	4.112	0	.038	.37	1.241	1.44

Note.—$df = 2$.
* $p = .1$.

The relation between the predictions of the constant-ratio model (Equation 15) and the similarity hypothesis (Equation 16) can be further investigated using the obtained estimates of the parameter d, reported in Table 2. It is easy to verify that the constant-ratio model is compatible with Equation 17 if and only if $d = 0$, while the similarity hypothesis implies $d > 0$. Hence, if the former holds, the estimates of d should be close to 0, whereas if the latter holds, the estimates should be substantially positive. (The magnitude of d should be interpreted in the light of the facts that all parameters are nonnegative and $c = 1$, see Equation 17 and Figure 7.) Inspection of Table 2 reveals that the majority of the d estimates in Task A are zero, while the majority of the d estimates in Tasks B and C are substantially positive. This agrees with the results of previous analyses (summarized in Table 1) according to which the constant-ratio model is satisfied in Task A, but not in Tasks B and C.

Taken together, the experimental findings suggest the hypothesis that the constant-ratio model is valid for choice among *unitary* alternatives (e.g., dots, colors, sounds) that are usually evaluated as wholes, but not for *composite* alternatives (e.g., gambles, applicants) that tend to be evaluated in terms of their attributes or components. This hypothesis is closely related to a suggestion made by Luce (1959):

> If we call a decision that is not subdivided into simpler decisions an elementary choice, then possibly we can hope to find Axiom 1 [Luce's choice axiom] directly confirmed for elementary choices but probably not for more complex ones [p. 133].

Research on multidimensional scaling based on similarity, or proximity, data (e.g., Shepard, 1964a; Torgerson, 1965) has also shown that judgments of unitary and composite stimuli (sometimes referred to as analyzable and unanalyzable) are governed by different rules. Much additional research, however, is required in order to assess the validity and the generality of the proposed hypothesis.

Finally, the distinction between unitary and composite stimuli is logically independent of whether the inconsistency reflected in choice probabilities is attributable to imperfect discrimination or to a conflict among incompatible criteria. (For a discussion of this last distinction, see Block & Marschak, 1960.) Although choice experiments in psychophysics typically involve imperfect discrimination with unitary stimuli while preference experiments are usually concerned with conflict among composite alternatives, the other two combinations also exist.

DISCUSSION

Strategic Implications

A major feature of the elimination-by-aspects model is that the probability of selecting an alternative depends not only on its overall value, but also on its relations to the other available alternatives. This gives rise to study of strategic factors in the design and the presentation of choice alternatives. Specifically, the present model provides a method for investigating questions concerning optimal design or location of alternatives in order to maximize (or minimize) choice probability under specified constraints. The following examples are intended to illustrate the scope and the nature of such a study.

First, consider a problem of binary comparisons. Suppose y and z are given, and we search for x such that $P(x;y)$ is maximized under the constraints that z has no aspects in common with any other alternative, and that $P(y;z)$ and $P(x;z)$ are fixed. By the former constraint, z can be viewed as a standard of comparison. Hence, the latter constraint can be interpreted as meaning that the overall values of y and of x (evaluated relative to z) are held fixed. Thus, only the position of x relative to y can be varied to maximize $P(x;y)$. Under these conditions, the present model implies that if $P(x;z) > P(y;z)$, x' should include as much of y' as possible. If, on the other hand, $P(x;z) < P(y;z)$, x' should include as little of y' as possible. The degree of overlap between x' and y' can be regarded as an index of the difficulty of comparing them. If x' includes y', the comparison is trivial, and $P(x;y)$ is maximal. If x' and y' are disjoint, the comparison is much more difficult, and $P(x;y)$ is less extreme.

In the light of this interpretation, the above result asserts that it is in the best interest of the favored alternative to make the comparison as easy as possible, while it is in the best interest of the nonfavored alternative to make the comparison as difficult as possible. This certainly makes sense: any increase in the difficulty of comparing the alternatives adds "error" to the judgment process and makes $P(x;y)$ closer to $1/2$. According to this logic, drastically different policies are prescribed depending on whether x is the favored or the nonfavored alternative. Advertising campaigns based on slogans such as "All aspirins are the same—why pay more?" and "This car is completely different from any other car in its class," illustrate, respectively, the policies recommended to the favored and the nonfavored alternatives. Note that these policies could be employed in the design of products as well as in their advertisements.

Second, let $T = \{x,y,\cdots,z\}$ and suppose that all pairwise choice probabilities are fixed and that we wish to select a set $A \subseteq T$ which includes both x and y so that the ratio $P(x,A)/P(y,A)$ is maximized. According to the elimination-by-aspects model, the above ratio is maximized when A consists of alternatives (which are not dominated by y) that "cover" as much of y as possible without "covering" much of x. If x and y are products in some market A, for example, then the present model predicts that the relative advantage of x over y is maximized when the other available products are as similar to y and dissimilar to x as possible. The example of choice among records discussed in the introduction and the similarity effect demonstrated in Table 1 illustrate the point. Note that this maximization problem cannot be investigated in Luce's model (Equation 2), for example, since by the constant-ratio rule $P(x,A)/P(y,A) = P(x;y)/P(y;x)$, $x, y \in A$, and hence is independent of A. According to the EBA model, in contrast, the above ratio can, in principle, be arbitrarily large, provided $P(x;y) \neq 0$.

Thus, if the present theory is valid, one can take advantage of the so-called "irrelevant alternatives" to influence choice probabilities. This result is based on the idea that the introduction of an additional alternative "hurts" similar alternatives more than dissimilar ones. This is a familiar notion in the context of group choice. The present development suggests that it is an important determinant of individual choice behavior as well. In practice, problems such as the design of a product or a political campaign involve many specific constraints concerning the nature of the product or the candidate. To the extent that these constraints can be translated into the present framework, the elimination-by-aspects model can be used (or abused) to determine the optimal design, or location, of choice alternatives.

Psychological Interpretation

The EBA model accounts for choice in terms of a covert elimination process based on sequential selection of aspects. Any such sequence of aspects can be regarded as a particular state of mind which leads to a unique choice. In light of this interpretation, the choice mechanism at any given moment in time is entirely deterministic; the probabilities merely reflect the fact that at different moments in time different states of mind (leading to different choices) may prevail. According to the present theory, choice probability is an increasing function of the values of the relevant aspects. Indeed, the elimination-by-aspects model is compensatory in nature despite the fact that at any given instant in time, the choice is assumed to follow a conjunctive (or a lexicographic) strategy. Thus, the present model is compensatory "globally" with respect to choice probability but not "locally" with regard to any particular state of mind.

In the proposed model, aspects are interpreted as desirable features; the selection of any particular aspect leads to elimination of all alternatives that do not contain the selected aspect. Following the present development, one can formulate a dual model where aspects are interpreted as disadvantages, or regrets, associated with the alternatives. According to such a model, the selection of a particular aspect leads to the elimination of all alternatives that contain the selected aspect. This model is also based on the notion of elimination by aspects, except that here an alternative is chosen if and only if none of its aspects is selected, whereas in the model developed in this paper an alternative is selected if and only if it includes all the selected aspects. The former model may be more appropriate when the defining features of the alternatives are naturally viewed as undesirable. In choosing among various insurance policies, for example, it may be more natural to apply the strategy of elimination by aspects to the various risks and premiums, treated as disadvantages or regrets, than to interpret them as relative advantages with respect to some reference points.[8]

Although the present model has been introduced and discussed in terms of aspects, we have shown that it requires no specific assumptions concerning the structure of these aspects. In the course of the investigation, however, assumptions concerning the structure and/or the relative weights of aspects were sometimes introduced. In discussing the Paris–Rome problem, for example, we assumed that Paris + (i.e., a trip to Paris plus an added bonus) includes Paris in the sense that all aspects of the latter are included in the former. Similarly, in analyzing Debreu's example, we assumed that the two recordings B_1 and B_2 of the Beethoven symphony are very similar to each other, whereas the suite by Debussy is relatively dissimilar to either of them. Essentially the same assumption was employed in the analysis of the experimental data. In all these instances, specific assumptions about the structure or the relative weights of aspects were added to the model on the basis of some prior analysis of the alternatives. The addition of such assumptions strengthens the predictions of the model and tightens its empirical interpretation. These assumptions, however, must be carefully examined because the inadequacy of an added assumption can erroneously be interpreted as a failure of the model.

---

[8] George Miller remarked that people seem to be better at finding what is wrong with an alternative than what is good about it. This certainly is true of some people, who might then find the "negative" version of the model less objectionable or more compatible with their way of thinking.

To illustrate this point, consider the following example of choice between articles of clothing. Let J denote a jacket, S a pair of matching slacks, and C a coat. Suppose that the coat is more valuable than the jacket, so $P(\mathrm{C};\mathrm{J}) > 1/2$. But since the slacks and the jacket are well matched, $P(\mathrm{JS};\mathrm{CS}) > 1/2$, where JS and CS denote the options consisting of the combined respective articles. Both JS and CS share the same article, S; hence one might be tempted to interpret S as a collection of aspects shared by the two alternatives. According to the elimination-by-aspects model, such aspects could be deleted without affecting the choice process. Consequently, under the proposed interpretation of S, $P(\mathrm{JS};\mathrm{CS}) = P(\mathrm{J};\mathrm{C})$ contrary to the assumptions. Further reflection, however, reveals that the interpretation of S as a collection of aspects common to both options is inappropriate. The fact that the jacket and the slacks form an attractive outfit implies that this alternative has some gestaltlike properties, or that the option JS includes some aspects that are not included in either J or S alone. Hence, the fact that the option JS includes both J and S as components does not, by itself, justify the conclusion that the aspects of JS can be partitioned into those associated with J and S alone.

Rational Choice and the Logic of Elimination by Aspects

The following television commercial serves to introduce the problem. "There are more than two dozen companies in the San Francisco area which offer training in computer programming." The announcer puts some two dozen eggs and one walnut on the table to represent the alternatives, and continues: "Let us examine the facts. How many of these schools have on-line computer facilities for training?" The announcer removes several eggs. "How many of these schools have placement services that would help find you a job?" The announcer removes some more eggs. "How many of these schools are approved for veterans' benefits?" This continues until the walnut alone remains. The announcer cracks the nutshell, which reveals the name of the company and concludes: "This is all you need to know in a nutshell."

This commercial illustrates the logic of elimination by aspects; it also suggests that this logic has some normative appeal as a method of choosing among many complex alternatives. The appeal of this logic stems primarily from the fact that it is easy to state, defend, and apply. In choosing among many complex alternatives such as new cars or job offers, one typically faces an overwhelming amount of relevant information. Optimal policies for choosing among such alternatives usually require involved computations based on the weights assigned to the various relevant factors, or on the compensation rates associated with the critical variables. Since man's intuitive computational facilities are quite limited (Shepard, 1964b; Slovic & Lichtenstein, 1971), the above method is difficult to apply.

Moreover, it seems that people are reluctant to accept the principle that (even very important) decisions should depend on computations based on subjective estimates of likelihoods or values in which the decision maker himself has only limited confidence. When faced with an important decision, people appear to search for an analysis of the situation and a compelling principle of choice which will resolve the decision problem by offering a clear-cut choice without relying on estimation of relative weights, or on numerical computations. (Altogether people seem to have more confidence in the rationality of their decisions than in the validity of their intuitive estimates, and the fact that the former depends on the latter is often met with a mixture of resistance and unhappiness.)

The strategy of elimination by aspects (illustrated by the above commercial) provides an example of such a principle: It is relatively easy to apply, it involves no numerical computations, and it is easy to explain and justify in terms of a priority ordering defined on the aspects. Inasmuch as people look for a decision rule that not only looks sensible, but which also seems easy to defend to oneself as well as to others, the principle of elimination by aspects appears attractive. Its uncritical application, however, may lead to very poor decisions. For virtually any available alternative, no matter how inadequate it might be, one can devise a sequence of selected aspects or, equivalently, describe a particular state of mind that leads to the choice of that alternative.

Indeed, the purpose of advertisement is to induce a state of mind in the decision maker which will result in the purchase of the advertised product. This is typically accomplished by increasing the salience and the availability of the desired state of mind. Being influenced by such factors, people are often lured into adopting a state of mind which, upon further reflection, appears atypical or inadequate. Shepard (1964b) tells of a person who is induced to purchase the *Encyclopedia Britannica* by imagining how he would read it in his free time and impress his friends with his newly acquired knowledge. Only after failing to consult the *Encyclopedia Britannica* for a long period of time does the person realize how inappropriate the state of mind was that had led him to purchase those many dusty volumes.

From a normative standpoint, the major flaw in the principle of elimination by aspects lies in its failure to ensure that the alternatives retained are, in fact, superior to those which are eliminated.

In the problem addressed by the above commercial, for instance, the existence of placement services that would help the trainee to find a job is certainly a desirable aspect of the advertised program. Its use as a criterion for elimination, however, may lead to the rejection of programs whose overall quality exceeds that of the advertised one despite the fact that they do not offer placement services.

In general, therefore, the strategy of elimination by aspects cannot be defended as a rational procedure of choice. On the other hand, there may be many contexts in which it provides a good approximation to much more complicated compensatory models and could thus serve as a useful simplification procedure. The conditions under which the approximation is adequate, and the manner in which this principle could be utilized to facilitate and improve decision making, are subjects for future investigations.

REFERENCES

BECKER, G. M., DEGROOT, M. H., & MARSCHAK, J. Stochastic models of choice behavior. *Behavioral Science*, 1963, **8**, 41–55. (a)

BECKER, G. M., DEGROOT, M. H., & MARSCHAK, J. Probabilities of choices among very similar objects. *Behavioral Science*, 1963, **8**, 306–311. (b)

BLOCK, H. D., & MARSCHAK, J. Random orderings and stochastic theories of responses. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, & H. Mann (Eds.), *Contributions to probability and statistics*. Stanford: Stanford University Press, 1960.

CHANDLER, J. P. STEPIT—Finds local minima of a smooth function of several parameters. *Behavioral Science*, 1969, **14**, 81–82.

CHIPMAN, J. S. Stochastic choice and subjective probability. In D. Willner (Ed.), *Decisions, values, and groups*. Vol. 1. New York: Pergamon Press, 1960.

COOMBS, C. H. On the use of inconsistency of preferences in psychological measurement. *Journal of Experimental Psychology*, 1958, **55**, 1–7.

COOMBS, C. H. *A theory of data*. New York: Wiley, 1964.

DEBREU, G. Review of R. D. Luce, Individual choice behavior: A theoretical analysis. *American Economic Review*, 1960, **50**, 186–188.

ESTES, W. K. A random-walk model for choice behavior. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959*. Stanford: Stanford University Press, 1960.

FISHBURN, P. C. Utility theory. *Management Science*, 1968, **13**, 435–453.

KRANTZ, D. H. *The scaling of small and large color differences*. (Doctoral dissertation, University of Pennsylvania) Ann Arbor, Mich.: University Microfilms, 1964. No. 65-5777.

KRANTZ, D. H. Rational distance function for multidimensional scaling. *Journal of Mathematical Psychology*, 1967, **4**, 226–245.

LANCASTER, K. J. A new approach to consumer theory. *Journal of Political Economy*, 1966, **74**, 132–157.

LUCE, R. D. *Individual choice behavior: A theoretical analysis.* New York: Wiley, 1959.

LUCE, R. D., & RAIFFA, H. *Games and decisions.* New York: Wiley, 1957.

LUCE, R. D., & SUPPES, P. Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology, III.* New York: Wiley, 1965.

MARSCHAK, J. Binary-choice constraints and random utility indicators. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford: Stanford University Press, 1960.

MORRISON, H. W. Testable conditions for triads of paired comparison choices. *Psychometrika*, 1963, **28**, 369–390.

RESTLE, F. *Psychology of judgment and choice.* New York: Wiley, 1961.

RUMELHART, D. L., & GREENO, J. G. Similarity between stimuli: An experimental test of the Luce and Restle choice models. *Journal of Mathematical Psychology*, 1971, **8**, 370–381.

SHEPARD, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, **1**, 54–87. (a)

SHEPARD, R. N. On the subjectively optimum selection among multiattribute alternatives. In M. W. Shelly & G. L. Bryan (Eds.), *Human judgments and optimality.* New York: Wiley, 1964. (b)

SLOVIC, P., & LICHTENSTEIN, S. C. Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 1971, **6**, 649–744.

THURSTONE, L. L. A law of comparative judgment. *Psychological Review*, 1927, **34**, 273–286.

THURSTONE, L. L. *The measurement of values.* Chicago: University of Chicago Press, 1959.

TORGERSON, W. S. Multidimensional scaling of similarity. *Psychometrika*, 1965, **30**, 379–393.

TVERSKY, A. Intransitivity of preferences. *Psychological Review*, 1969, **76**, 31–48.

TVERSKY, A. Choice by elimination. *Journal of Mathematical Psychology*, 1972, in press.

TVERSKY, A., & RUSSO, J. E. Similarity and substitutability in binary choices. *Journal of Mathematical Psychology*, 1969, **6**, 1–12.

APPENDIX

This appendix estalishes the equivalence of the two formulations (Equations 6 and 9) of the elimination-by-aspects model. Let T be a finite set of alternatives. For each $x \in T$, let x' denote the set of aspects associated with x. For any $A \subseteq T$, define $A' = \{\alpha | \alpha \in x'$ for some $x \in A\}$, $A^0 = \{\alpha | \alpha \in x'$ for all $x \in A\}$, and $\bar{A} = \{\alpha | \alpha \in x'$ for all $x \in A$ & $\alpha \notin y$ for any $y \notin A\}$. We wish to show that there exists a positive scale u on $T' - T^0$ satisfying Equation 6 if and only if there exists a scale U on $\{\bar{A}_i | A_i \subset T\}$ satisfying Equation 9.

It follows at once from the above definitions that $\{\bar{A}_i | A_i \subset T\}$ forms a partition of $T' - T^0$, since any $\alpha \in T' - T^0$ belongs to exactly one $\bar{A}_i$. Suppose Equation 6 holds. For any $A \subset T$, define $U(\bar{A}) = \sum_{\alpha \in \bar{A}} u(\alpha)$. By the positivity of u, U is nonnegative and $U(\bar{A}) = 0$ iff $\bar{A} = \phi$. Note that if $\alpha, \beta \in \bar{B}$ then for all $A \subset T$, $A_\alpha = A_\beta = A \cap B$. Furthermore, since $\{\bar{B}_i | x \in B_i\}$ forms a partition of x', the numerator in Equation 6 can be expressed as

$$\begin{aligned}\sum_{\alpha \in x' - A^0} & P(x,A_\alpha)u(\alpha) \\ &= \sum_{x \in B_i \not\supseteq A} \sum_{\alpha \in \bar{B}_i} P(x,A_\alpha)u(\alpha) \\ &= \sum_{x \in B_i \not\supseteq A} P(x,A \cap B_i) \sum_{\alpha \in \bar{B}_i} u(\alpha) \\ &= \sum_{B_i \not\supseteq A} P(x,A \cap B_i)U(\bar{B}_i).\end{aligned}$$

(The condition $x \in B_i$ under the summation sign is deleted because for any $x \notin B_i$, $P(x, A \cap B_i) = 0$. Similarly, since $\{\beta | \beta \in A' - A^0\} = \{\beta | \beta \in \bar{A}_j$ for some A_j such that $A_j \not\supseteq A$ and $A_j \cap A \neq \phi\}$, the denominator in Equation 6 can be expressed as

$$\sum_{\beta \in A' - A^0} u(\beta) = \sum_{A_j \in \mathfrak{A}} U(\bar{A}_j)$$

where

$$\mathfrak{A} = \{A_j | A_j \cap A \neq A, \phi\}.$$

Thus, Equation 6 reduces to Equation 9, since

$$\frac{\sum_{\alpha \in x' - A^0} u(\alpha)P(x,A_\alpha)}{\sum_{\beta \in A' - A^0} u(\beta)} = \frac{\sum_{B_i \not\supseteq A} U(\bar{B}_i)P(x, A \cap B_i)}{\sum_{A_j \in \mathfrak{A}} U(\bar{A}_j)}$$

Conversely, suppose Equation 9 holds. That is, there exists a scale U such that $P(x,A)$ is given by the right-hand side of the above equation. For any $x \in T$, let $x' = \{A_i \subset T | x \in A_i\}$, and $u = U$, hence Equation 9 reduces to Equation 6. Finally, if either of the above denominators vanishes, so does the other.

(Received February 7, 1972)

Psychological Review
1972, Vol. 79, No. 4, 300–316

ABSOLUTE JUDGMENT AND PAIRED-ASSOCIATE LEARNING:

KISSING COUSINS OR IDENTICAL TWINS?[1]

JANE A. SIEGEL[2] AND WILLIAM SIEGEL[3]

University of Michigan

The method of absolute judgment is examined in parallel with the paired-associate learning paradigm, to which it bears a striking procedural resemblance. It is shown that a number of the traditional learning variables, including stimulus meaningfulness, the retention interval, ordinal position, and practice, influence performance in absolute judgment. Moreover, stimulus information is shown to be confounded with the average retention interval, and the concept of a "channel capacity" for judgment is evaluated in light of this finding. It is concluded that absolute judgment performance is limited by the memory span for sensory stimuli, and it may be greatly improved by organization or chunking.

The method of paired-associate learning has for many years been a standard paradigm in the field of verbal learning and memory. The absolute judgment task has also had a long history, but within the area of psychophysics. As Wever and Zener (1928) have pointed out, the absolute judgment method had its origins in the work of G. T. Fechner, L. Witmer, D. R. Major, and J. Cohn in the nineteenth century, when it emerged as a variant of the method of constant stimuli. More recently, the absolute judgment method has been used to estimate the amount of information which can be transmitted by the human observer (e.g., Attneave, 1959; Garner & Hake, 1951; Miller, 1956).

Because absolute judgment and paired-associate learning represent different traditions within psychology, each with its own literature, it is easy to assume that the tasks are procedurally quite different and that they tap different psychological processes as well. However, it is our contention that this assumption is incorrect. We began to explore this possibility after noticing a striking formal similarity between the two paradigms, a similarity which has also been noted by such writers as Gibson (1953), Garner (1962), Bush, Galanter, and Luce (1963), and Postman (1963). A fine-grained examination of the data of these two tasks has led us to the conclusion that the resemblance is more than skin-deep, and that common psychological processes underlie both paradigms.

METHODOLOGICAL COMPARISON

Wever and Zener (1928) presented a detailed review of the absolute judgment literature, and it is clear that at that time the absolute judgment paradigm was equivalent to what would now be called a category-rating task. The subject used numerical categories provided by the experimenter to report the magnitude of his sensory experience. On any trial, he was free to rate or judge the stimulus using whichever response he felt to be most appropriate.

The introduction of information theory was associated with two important changes in the absolute judgment method. First, the investigator determined a "correct" mapping of responses onto stimuli, and the subject was required to respond as *accurately* as possible. Second, the experimenter often informed subjects of the correct response after each trial. These procedural modifications changed the emphasis in absolute judgment from judgment to identification.[4] As a result, the absolute judgment paradigm emerged with a basic formal similarity to another identification task, the anticipation method of paired-associate learning used in human-learning studies.

---

[1] Jane Siegel was supported by Predoctoral Fellowship No. 1F01 MH 42833-01 from the National Institute of Mental Health, and William Siegel was supported by a Rackham Fellowship from the University of Michigan. A number of individuals were kind enough to read and criticize several earlier drafts of this paper. In particular, we thank Arthur Melton, Allan Paivio, Irwin Pollack, Richard Pew, and Endel Tulving for their comments. To Keith Smith, Daniel Weintraub, and Robert Zajonc we owe many hours of valuable discussion, and we are glad to acknowledge their encouragement. Finally, we would like to thank Academic Press and the Acoustical Society of America for permission to reproduce several of the figures included in this paper.

[2] Requests for reprints should be sent to Jane Siegel, who is now at the Department of Psychology, University of Western Ontario, London, Ontario, Canada.

[3] Now at the University of Western Ontario.

Both the absolute judgment and paired-associate learning methods involve an arbitrary pairing of response with each member of the stimulus set. Following this initial pairing, the stimuli are presented in random order and the subject is required to produce the response appropriate to each one. Of course there are typical differences between absolute judgment and paired-associate learning, but enough exceptions exist so as to blur the distinction between the two paradigms.

We usually think of absolute judgment as involving "sensory" stimuli and paired-associate learning as using "verbal" stimuli. In the absolute judgment task, the stimuli are ordered along one or more continua, and the responses are ordered as well. Moreover, responses are mapped onto stimuli so that the two orders coincide. This is done so that the subject can "judge," that is, base the magnitude of his response on a sensory magnitude. In paired-associate learning, both stimuli and responses usually consist of unordered words or nonsense syllables. However, it is easy to demonstrate that stimulus–response ordering does not provide a critical distinction between the two tasks. For example, the following ordered stimuli have been used in paired-associate learning: (*a*) weights and time intervals (Murdock, 1960); (*b*) circles varying in diameter (McGuire, 1961); (*c*) differences in the percentage reflectance of gray (Ebenholtz, 1966); (*d*) spatial positions (Jensen, 1962); (*e*) a two-dimensional array of lights (Young & Buck, 1968); and (*f*) the number series (Ebenholtz, 1966). Alphanumeric characters are commonly used as responses in paired-associate learning as well as in absolute judgment. In order to test his model of paired-associate learning, McGuire (1961) used one condition in which the stimuli were nine circles varying in diameter, the responses were the numbers 1–9, and the stimulus and response orders coincided. Perhaps it is more meaningful to consider stimulus–response ordering as factors which are important in an *identification* task, rather than as defining characteristics of absolute judgment and paired-associate learning.

The remaining differences between the two paradigms seem to reflect attitudes and biases associated with verbal learning and psychophysics. The task of the subject in paired-associate learning is to acquire associations between stimuli and responses. Thus conditions of stimulus–response pairing and feedback are carefully controlled. In absolute judgment, the goal of the subject is to judge the stimuli presented using the system of responses indicated by the experimenter. Typically he is given some initial practice to familiarize himself with the task and then is exposed to a long series of test trials, often without feedback. In a series of experiments involving pitch judgments, Pollack (1952, 1953) allowed extensive practice, and feedback was given throughout. However, he made it clear that practice was designed to improve judgments rather than associations between stimuli and responses.

The basic data of both absolute judgment and paired-associate learning consist of a sequence of categorical responses made to the stimulus series. Depending on the viewpoint of the investigator, however, the same raw data may be analyzed and interpreted quite differently. Since in paired-associate learning at least part of the purpose of the experiment is to chart the course of acquisition, performance is usually plotted as a function of trials, and the experiment is designed so as not to confound the effects of other manipulated variables with those of practice. In absolute judgment, it is tacitly assumed that learning has reached asymptote, and usually no attempt is made to assess the effects of practice. To the psychophysicist, a confusion reflects properties inherent in the stimulus which prevent discrimination. The learning theorist, on the other hand, may regard almost any stimulus as a possible cue for discrimination. To him, a confusion is an indication that learning has not yet reached asymptote.

---

[4] As Bush et al. (1963) made clear, there are a number of absolute judgment tasks that are commonly used in psychology, including category scaling and magnitude estimation. In this paper, we will restrict the discussion of "absolute judgments" to identification tasks where the experimenter specifies a one-to-one mapping of responses onto stimuli.

To summarize, in both absolute judgment and paired-associate learning the subject is presented with the members of a stimulus set in random order, and he is required to produce the response which the experimenter has paired with each item. Ordered stimuli and responses are more common in judgment than in paired-associate learning experiments, but the use of ordered material is not exclusive to absolute judgment. A major difference between the two paradigms concerns the manner in which the raw data are analyzed and interpreted by the experimenter. This difference is more a matter of theoretical viewpoint than of substance, and we have been unable to specify any major procedural distinctions between absolute judgment and paired-associate learning.

McGUIRE'S MULTIPROCESS MODEL OF PAIRED-ASSOCIATE LEARNING AND THE ABSOLUTE JUDGMENT CHANNEL CAPACITY

Although absolute judgment and paired-associate learning are nearly identical procedurally, there is a striking difference in the results that are typically obtained in the two tasks. In paired-associate learning, performance improves regularly as a function of trials, and if sufficient practice is allowed, subjects can learn to identify large numbers of items. In absolute judgment, performance also improves with practice (e.g., Hartman, 1954), but learning effects are usually considered to be unimportant. Even after a large number of trials with feedback, most observers cannot identify more than five or seven items without error. This finding has led a number of reviewers (e.g., Attneave, 1959; Garner, 1962; Miller, 1956) to conclude that there is a fixed upper limit or channel capacity on absolute judgment performance. However, the psychological mechanisms which underlie this limitation have not been clearly specified.

Considering the similarity between absolute judgment and paired-associate learning, it seems fitting to use McGuire's (1961) multiprocess model of paired-associate learning as a heuristic tool for pinpointing the locus of subjects' difficulty in absolute judgment. McGuire's hypothesis is that paired-associate learning is distinguished by three stages: (*a*) a *stimulus-learning* phase, in which the subject establishes a mediating response to each member of the stimulus set; (*b*) a *response-learning phase,* which involves integration of members of the response set; and (*c*) a *hookup* phase, in which the overt paired-associate response is associated with the mediating response to the stimulus. It is perhaps fitting that McGuire tested his model of paired-associate learning using materials which are more often employed in absolute judgment. His stimuli were nine circles varying in diameter, and the responses were numbers.

In the absolute judgment task, the responses usually consist of single-digit or two-digit numbers, and so response learning should not present any difficulty for the subject. When the same responses are used in paired-associate learning with verbal stimuli, learning appears to progress rapidly and nothing like the channel capacity is encountered. Furthermore, the hookup phase should not be a problem in absolute judgment, since the responses are ordered in the same manner as the stimuli. Thus it appears to us that the locus of the subjects' difficulty in absolute judgment is the stimulus-learning phase. The question remains as to why stimulus learning provides such an obstacle in absolute judgment.

Dimensionality versus Meaningfulness as Determinants of the Channel Capacity

The most typical procedural difference between absolute judgment and paired-associate learning is that unidimensional sensory stimuli are used in the former and multidimensional verbal stimuli are used in the latter. Thus, it seems reasonable to attribute the different performance levels characteristic of these tasks to the dimensionality of the stimuli used in them. In fact, the information theorists have suggested that unidimensional stimuli are perceptually impoverished, and a number of studies have shown that information transmission is greatly improved when the dimensionality of the display is increased (Eriksen & Hake, 1955; Klemmer & Frick, 1953; Pollack & Ficks, 1954).

While stimulus dimensionality appears to be a critical variable in an identification task, the possibility remains that some other factor covaries with dimensionality and is responsible for at least part of its effect. In fact, there are a number of instances where stimuli are identified more or less easily than the number of dimensions on which they vary would indicate. For example, stimuli at the ends of a unidimensional continuum are usually judged quite accurately (e.g., Alluisi & Sidorsky, 1958; Murdock, 1960). Some sets of unidimensional stimuli yield larger-than-average values of information transmission. The channel capacity for angle of inclination is 4.5 bits (Muller, Sidorsky, Slivinske, Alluisi, & Fitts, 1955), which corresponds to the perfect identification of about 23 items. Most subjects can identify 11–13 hues (Chapanis & Halsey, 1956; Conover, 1959), but only 3–5 items varying in brightness (Eriksen & Hake, 1955), loudness (Garner, 1953), or pitch (Hartman, 1954; Pollack, 1952, 1953; W. Siegel, 1972). With the pitch continuum, there are large individual differences, and some subjects appear to be able to identify many tone frequencies without error. Attneave (1959) cites an unpublished study in which an individual with "perfect pitch," the concert master of a symphony orchestra, was capable of transmitting 5.5 bits, a level equivalent to the perfect identification of 45 tones.

The identification of complex, multidimensional stimuli is not always as easy or immediate as one might suspect. Corcoran, Carpenter, Webster, and Woodhead (1968) attempted to teach eight groups of subjects to identify 14 simulated engine noises, using a variety of training procedures. All groups showed continuous improvement with practice, but even on the tenth and final session, no group exceeded 50% correct. Similarly Webster, Carpenter, and Woodhead (1968) manipulated the harmonics present in buzz-type tonal patterns, using both discrimination and identification tasks. Nine complex sounds were presented, and the subjects were quite accurate at discriminating differences between them. However, even though there were 12 identification sessions, the percentage of correct responses did not reach 40% on any one of them. This level of performance is on a par with the lowest channel capacities found for unidimensional continua, and is substantially worse than the accuracy of identification of hue and of angle of inclination.

It is clear that stimulus dimensionality is far from perfect as a predictor of identification performance. We feel that there is a variable which is at least partially correlated with dimensionality, and which is responsible for much of the variance in absolute judgment and paired-associate learning. That variable is stimulus *meaningfulness,* a factor which has been studied extensively in the learning literature. Traditionally, meaningfulness has been defined as the percentage of subjects who respond with a verbal association to a stimulus during a fixed time period (Archer, 1960; Glaze, 1928; Hull, 1933; Witmer, 1935). Meaningfulness has also been operationalized in terms of the *number* of associations given to an item during a standard time period (Noble, 1952), and in terms of *rated* association value (Noble, Stockwell, & Pryor, 1957).

It is often found that high-meaningful stimuli produce more rapid learning in paired-associate learning than low-meaningful items. According to Martin (1968), the greater effectiveness of high-meaningful stimuli is not the result of their ability to elicit a large number of associations, as traditional measures imply; rather, Martin feels that the critical factor is that high-meaningful items tend to elicit the same association from trial to trial, that is, variability of encoding is low for such items. Paivio (1971) has also questioned the traditional definitions of meaningfulness, and he has developed a theory of meaning in which an item's ability to evoke an imaginal code plays a central role. Moreover, Paivio has shown that the imagery dimension is highly correlated with traditional meaningfulness measures, and it appears to account for most of the effect of stimulus meaningfulness in paired-associate learning.

Regardless of which definition of meaning one prefers, it is evident that the sensory stimuli usually used in absolute judgment would fall at the extreme low end of the continuum. Unidimensional stimuli are seldom encountered outside of the laboratory, and subjects seem to have few overlearned associations, either verbal or imaginal, to simple sensory stimuli. In fact, using any of the above definitions, most unidimensional stimuli would approach zero meaningfulness. In contrast, all of the stimuli typically used in verbal-learning studies are relatively high-meaningful stimuli. A letter, trigram, or word can invariably elicit at least one reliable associate: the verbalization or pronunciation of the stimulus item. Because of the ubiquity of names and labels with verbal stimuli, these associations have largely been ignored, but with sensory stimuli, the availability of names and labels is not to be taken for granted.

While most unidimensional stimuli may be characterized as extremely low in meaningfulness, there are some sensory continua which may be reliably labeled. Most of us can name at least the vertical and horizontal when judging angle of inclination. There are distributions of color names associated with hues (Beare, 1963; Lenneberg 1957), and these distributions depend on the culture one grows up in (Lenneberg, 1961). Individuals with "perfect pitch" use note names corresponding to the diatonic scale to identify tone frequencies.

When prior associations are available to the subject, an identification task does not involve original learning, but rather transfer of training. Associations which are already present can mediate the acquisition of new responses. The existence of mediational processes in absolute judgment is indicated by "bumps" in the performance curve where certain stimuli are judged with high accuracy. "Bumps" have been observed in the judgment of hue (Eriksen & Hake, 1957; Volkman & Engen, 1961), angle of inclination (Muller et al., 1955), and linear interpolation (Levett, 1952; Nash, 1964; Spitz, 1967), but are not found in most other continua. These "bumps" seem to represent stimuli which are namable, such as the vertical and horizontal, and the middle of a line. It is interesting that the unidimensional continua which are associated with names and labels also have the highest channel capacities, while the multidimensional stimuli which are the most difficult to judge or identify seem to lack consistent verbal associations. It is our hypothesis, then, that meaningfulness is a better predictor of absolute judgment or paired-associate learning performance than dimensionality, and that most sensory stimuli are at the low end of the meaningfulness continuum.

Meaningfulness and the Retention Interval

Our characterization of absolute judgment stimuli as ultra low meaningful gains in heuristic value if we can specify in more detail the mechanisms by which meaningfulness influences identification performance. An experiment by Martin and Melton (1970) demonstrates that with nonsense syllables as stimuli, meaningfulness influences the ease with which items are retained in memory. In their study, 160 trigrams were presented twice each in a continuous recognition task. Three levels of meaningfulness were employed in a between-groups design. On each trial, the subject indicated whether the trigram was "old" or "new," and successive occurrences of each stimulus were separated by 1, 3, 6, 15, or 30 intervening recognition trials. The results of the Martin-Melton study are reproduced in Figure 1. Here, performance accuracy corrected for false positives is plotted as a function of the retention interval, with meaningfulness as the parameter. At brief retention intervals, meaningfulness appeared to have little effect on recognition accuracy. But as the number of intervening items was increased, the rate of forgetting was inversely proportional to meaningfulness.

If our hypothesis is correct, and sensory stimuli generally approach zero meaningfulness, then we may extrapolate from the Martin-Melton data and predict that such items are subject to rapid forgetting. A number of recent studies using relative-judgment procedures suggest that this is the case. The relative-judgment task becomes a recognition-memory paradigm if there is a time delay between the standard and comparison stimuli, for then the subject must base his decision on the difference between the perceived value of the comparison and the remembered value of the standard. The time between the standard and comparison stimuli has been varied in studies investigating short-term memory for pitch (Massaro, 1970; Wickelgren, 1966, 1969), loudness (Kinchla & Smyzer, 1967; Tanner, 1961), visual movement (Kinchla & Allan, 1970), visual position (Allan, 1968), and kinesthetic position (Johnson, Chatelier, & Mocharnuk, 1969). In all of these experiments, the observers' sensitivity declined rapidly during the course of the retention interval. Moreover, the rate of forgetting was greatest during the first few seconds after the offset of the standard.

We have predicted that low-meaningful sensory stimuli are difficult to store effectively in memory, and this prediction receives strong support from the available data. It remains to be demonstrated what implications this finding has for performance in an identification task. Martin (1967a, 1967b) has demonstrated an important relationship between recognition memory and performance in paired-associate learning. Using the study–test method, Martin required his subjects to identify the stimulus on each trial as "old" or "new" before making the paired-associate response. He found that on those trials on which the subject failed to recognize the stimulus term, the probability of a correct association fell to chance level. From Martin's data, it is clear that stimulus recognition is necessary for correct association.

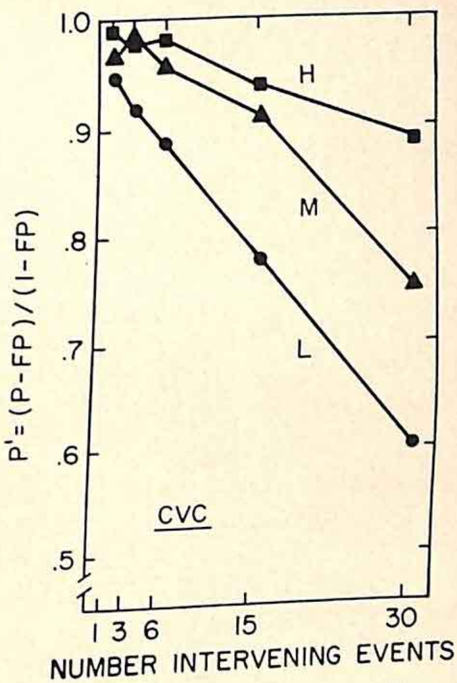


FIG. 1. Proportion of correct recognitions corrected for false positives (FP) as a function of number of intervening items and meaningfulness. (Reprinted with permission from an article by E. Martin and A. W. Melton published in the *Journal of Verbal Learning and Verbal Behavior,* 1970, Vol. 9. Copyrighted by Academic Press, 1970.)

Since recognition performance is dependent on the retention interval, one would expect to find short-term memory effects in an identification task. Peterson, Saltzman, Hillner, and Land (1962), Atkinson and Shiffrin (1968), and Calfee (1968) are among those who have shown that performance in paired-associate learning is an inverse function of the retention interval. In an identification task, the retention interval is determined by the time between successive presentations of a given pair. Other stimulus–response pairs occur during this interval, and thus time covaries with the number of intervening items.

In the absolute judgment paradigm as well as in paired-associate learning, the stimuli are presented in a random sequence, and the retention interval is an uncontrolled variable. On some trials, a stimulus repeats itself, and the subject must store information in memory only during the intertrial interval. More often, however, the number of intervening

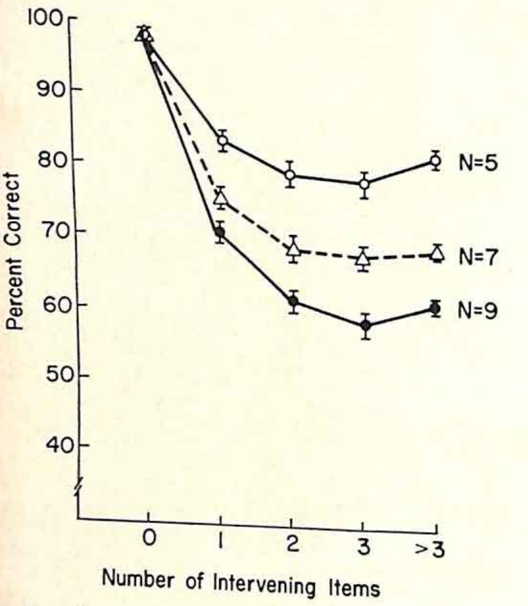


FIG. 2. Joint effects of the retention interval and the number of alternatives in the judgment of pitch. (Data are from an article by W. Siegel published in the *Journal of Experimental Psychology*, 1972, Vol. 94. Copyrighted by the American Psychological Association, 1972.)

items is greater than zero. Here, the subject must store information over a correspondingly greater period of time, and with low-meaningful sensory stimuli we might expect substantial forgetting to occur.

The role of the retention interval in absolute judgment was recently examined in a series of four studies (W. Siegel, 1972), and it was found that all subjects in all experiments showed rapid forgetting. Some typical data are shown in Figure 2. In this study, six observers judged sets of five, seven, and nine tone frequencies. The details of the experiment, such as the frequency range and mode of feedback, were similar to those of previous absolute judgment pitch studies (e.g., Pollack, 1952), except that performance was examined separately for each retention interval. Figure 2 demonstrates that at the briefest retention interval, subjects seldom made any errors, but after only one or two intervening items, accuracy was reduced to its lowest level. This rapid drop to asymptote is quite unlike the more protracted intervening-items effect observed in paired-associate learning with verbal stimuli (e.g., Brelsford, Keller, Shiffrin, & Atkinson, 1966). Forgetting in absolute judgment was found to result from both the passage of time (temporal decay) and from the interfering effects of listening to and judging other tones during the retention interval. Furthermore, when subjects were not required to store pitch information over time in absolute judgment, performance was as precise as that found in "discrimination" tasks.

We have characterized the simple sensory stimuli typically used in absolute judgment as ultra low meaningful, and this led to the prediction that they should be subject to rapid forgetting. The demonstration of strong retention interval effects in absolute judgment seems to confirm this hypothesis and provides an explanation for the low level of performance characteristic of this task. While subjects were capable of great accuracy in absolute judgment on repetition trials where the retention interval was brief, such repetition trials occur relatively infrequently, depending of course on the number of items being judged. Thus, overall absolute judgment performance primarily reflects the subjects' ability to differentiate stimuli under conditions of high memory load, and under these conditions, even relative judgment performance with sensory stimuli is greatly impaired.

Channel Capacity Reconsidered

So far we have limited our discussion of retention interval effects in absolute judgment to one dependent variable, response accuracy. However, it is common practice to apply the tools of information theory to absolute judgment data and to estimate the amount of information transmitted (H_t) in the subjects' responses. It is when this is done that the channel capacity phenomenon becomes apparent. When H_t values are plotted as a function of stimulus information (H_s), transmission scores first increase at a maximal rate and then remain invariant as H_s is made large. The asymptote on H_t appears to be in the neighborhood of 2.3 bits under optimal conditions in pitch judgment studies (Hartman, 1954; Pollack, 1952, 1953), and this is equivalent to the accurate identification of only five items.

An attractive feature of the information theoretic approach is that H_t values appear to remain relatively constant over changes in a number of experimental variables, such as the range and spacing of stimuli, practice, and feedback (Garner, 1962; Miller, 1956). Given the large effects of the retention interval on absolute judgment accuracy scores (W. Siegel, 1972), and given that the memory load varies from trial to trial in absolute judgment, it would be interesting to see if H_t is uninfluenced by retention effects.

In fact, the average retention interval in absolute judgment is confounded with the H_s measure of information theory. Since the absolute judgment experiment usually consists of a random sequence of equiprobable stimuli, the probability of an item repeating itself after X intervening trials is equal to $(1/N)(1-1/N)^X$, where N is the number of stimulus alternatives. The mean number of intervening items is simply $N-1$. Under these conditions, $H_s = \log_2 N$, and thus H_s is monotonically related to the average retention interval.

While it is true that H_s is confounded with the mean number of intervening items, it is still possible to manipulate H_s while holding the retention interval constant. This was done (W. Siegel, 1972) by generating separate confusion matrices and transmission scores for each number of intervening items, and the results are shown in Figure 3. It may be seen that overall performance, as represented by the heavy line, remains approximately constant as H_s varied from 2.32 to 3.17 bits. This is the same result which has been found in many other absolute judgment studies and which has been interpreted as evidence for a channel capacity. However, when transmission is examined separately for each retention interval, the invariance disappears, and at short retention intervals, H_t increases rapidly as H_s is made larger. The most interesting finding is that at the briefest retention interval used in this study, H_t values approached maximum for all levels of H_s. Only at longer retention intervals was the classic limitation on information transmission observed. Thus the absolute judgment channel capacity appears to be a limitation on short-term memory for sensory stimuli.

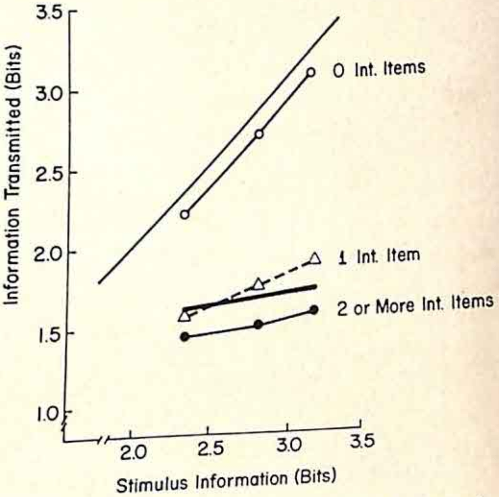


FIG. 3. Joint effects of stimulus information and the retention interval in the judgment of pitch. (Reprinted from an article by W. Siegel published in the *Journal of Experimental Psychology*, 1972, Vol. 94. Copyrighted by the American Psychological Association, 1972.)

END EFFECTS IN ABSOLUTE JUDGMENT AND PAIRED-ASSOCIATE LEARNING

Up to this point, we have discussed the role of such variables as stimulus dimensionality, meaningfulness and the retention interval in absolute judgment and paired-associate learning. It is our contention that by viewing these two tasks in parallel, one may observe similar empirical phenomena, and in addition it may be possible to reinterpret some of the major findings. Now we will consider another phenomenon which is common to both absolute judgment and paired-associate learning, the ubiquitous inverted-U-shaped error distributions, or *end effects*.

In the absolute judgment experiment, not all members of the stimulus set are judged with equal precision. Stimuli at the ends of the continuum are identified more accurately than those in the middle. This "end-anchoring" effect, as it is referred to in the absolute judgment literature, appears to be a genuine context effect, since the same physical stimuli are confused much more often when they are placed in the middle of the stimulus range than when they are end items (W. Siegel, 1972).

Although the terminology used to depict them is different, end effects appear to pervade the field of verbal learning as well. For example, the serial-position effect (Hovland, 1938; Young, 1968) refers to the fact that in serial learning studies, the number of errors made is an inverted-U-shaped function of an item's temporal position in the list. Although they are not as well known, it is more germane to the present discussion that a variety of end effects, or "ordinal position effects," have been reported in paired-associate learning as well as in serial learning. When the stimuli are ordered along sensory (Murdock, 1960), numerical (Ebenholtz, 1966), spatial (Jensen, 1962), or temporal (Rubin & Brown, 1967) continua, bow-shaped error curves are found. In the absolute judgment paradigm, both the stimulus and response ordering contribute to the observed end effects. In paired-associate learning, however, stimulus and response ordering have been varied independently. Data of Ebenholtz (1966) suggest that stimulus ordering is sufficient to produce ordinal position effects, but it also seems that bowed error distributions may occur when an ordered response set is mapped onto an unordered stimulus set (DeSoto & Bosley, 1962; Pollio, 1968; Pollio & Draper, 1966).

A number of theories have been put forward in both the perception and learning fields which attempt to account for ordinal position effects. The most simple of these is that subjects are biased toward the end responses, using them more frequently than the probability of occurrence of the end stimuli would dictate. This interpretation was ruled out (W. Siegel, 1972) by demonstrating that the usual end effects occur in absolute judgment even when the data are bias free.

A second plausible theory is that the end stimuli are confused less often because generalization can occur in only one direction at the ends, and in two directions in the middle. This view is implicit in a number of theories of judgment (e.g., Thurstone, 1927; Torgerson, 1958). Although this "unidirectional error" hypothesis is simple and attractive, an experiment by Eriksen and Hake (1957) suggests that it is also incorrect. These authors used a series of 20 hues in which the end points were not clearly defined. Since extreme violet is more similar to extreme red than either of these hues are to, say, yellow, one can think of the hue continuum as forming a closed circle. The unidirectional error hypothesis would not predict a bowed error distribution in this case, but Eriksen and Hake found the usual ordinal position effect. However, in their experiment, the most accurate performance was found with stimuli at the ends of the *response* continuum, that is, with whichever stimuli were labeled 1 and 20.

Eriksen and Hake have put forward a theory of their own in which memory processes play an important part, and which they have named the "subjective standard" hypothesis. According to this theory, some members of the stimulus set are stored in memory and used as standards for judging the remaining items. Typically, the subject selects the ends of the stimulus continuum for use as subjective standards. However, when the set of stimuli lacks distinctive end points, the subject may adopt a different mnemonic strategy.

It is interesting that Feigenbaum and Simon (1962) have proposed a theory of the serial position effect in verbal learning which is somewhat similar to the subjective standard hypothesis. Feigenbaum and Simon argue that since the subject is not capable of processing all of the items in a serial list at the start of training, he picks the most distinctive items, usually those at the ends of the list, as points at which to begin his learning task. Although it was developed for serial learning, the theory might easily apply to paired-associate-learning tasks using ordered material, and to the absolute judgment situation as well. Both this theory and that of Eriksen and Hake imply that the subject adopts selective rehearsal strategies based on the saliency of the stimulus materials, and both postulate memory mechanisms to account for end effects. If this line of argument is correct, then when memory factors are made minimal, the end effects should disappear.

An attempt was made to test this hypothesis (W. Siegel, 1972) by examining absolute judgment accuracy under conditions where no temporal retention was required. An experiment was conducted where subjects judged sets of five, seven, and nine tone frequencies with feedback. As is always the case in absolute judgment, the retention interval varied from trial to trial, but here the minimum retention interval was zero (0 intervening items/.00 second). Performance accuracy at each ordinal position is shown in Figure 4 for each of the three stimulus sets. As expected, U-shaped accuracy curves were found on nonrepetition trials where the subjects were forced to store pitch information in memory. But on repetition trials, where no temporal retention was required, the end effects were just as large as on nonrepetitions. This finding was entirely unexpected, and it provides strong evidence against a memory theory of the ordinal position effect, such as Eriksen and Hake's subjective standard hypothesis.

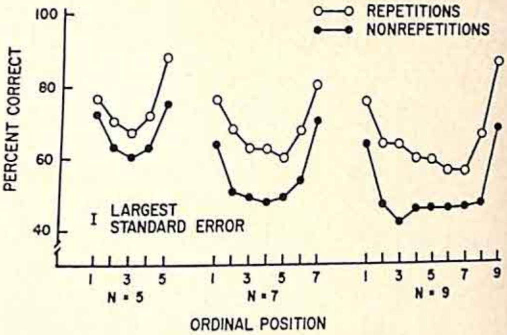


FIG. 4. Joint effects of the retention interval and ordinal position in the judgment of pitch. (Data are shown separately for repetition trials—0 intervening items/.00 second—and nonrepetition trials.) (Reprinted from an article by W. Siegel published in the *Journal of Experimental Psychology*, 1972, Vol. 94. Copyrighted by the American Psychological Association, 1972.)

It is most puzzling to us that subjects show typical end effects under conditions of zero memory load in an identification task. Perhaps the "unidirectional error" hypothesis, which lost favor as a result of the work of Eriksen and Hake, should be subjected to further experimental test. Another possibility is that the ordinal position effect is a genuine perceptual phenomenon representing an attentional strategy on the part of subjects. At the moment, though, one can only conclude that more research is needed to clarify the nature of the pervasive and elusive end effects in absolute judgment and paired-associate learning.

ORGANIZATION, CHUNKING, AND STIMULUS IDENTIFICATION

Earlier in this paper, we described McGuire's three-stage model of paired-associate learning, and used that theory as a heuristic device for pinpointing the locus of the subjects' difficulty in the absolute judgment task. As a model of paired-associate learning, however, McGuire's schema is too simple because it fails to make explicit the role of cognitive and organizational strategies which subjects seem to adopt when they are asked to learn a paired-associate list (cf. Battig, 1968; Mandler, 1967). Subjects routinely search for some rule or coding strategy which will allow them to generate the correct response upon presentation of the stimulus (e.g., Prytulak, 1971). Moreover, there is also evidence that rules are generated to structure the list into meaningful subsets of items. For example, Brown and Battig (1966) have shown that paired-associate learning lists may be grouped into sets of learned versus unlearned items, while Segal and Mandler (1967) have demonstrated that subjects may organize items in paired-associate learning on the basis of whether they are stimulus or response terms. According to Miller (1956) and Mandler (1967), organization or "chunking" allows one to overcome the initial limitation in any learning task imposed by the span of immediate memory.

While organization and chunking have seldom been referred to as such in the absolute judgment literature, there seems little doubt that these phenomena may be important there as well as in paired-associate learning. When the stimuli are spaced along a single continuum, then the obvious means of organizing the stimulus set is to partition the continuum into segments. This may be accomplished in several ways, one of which

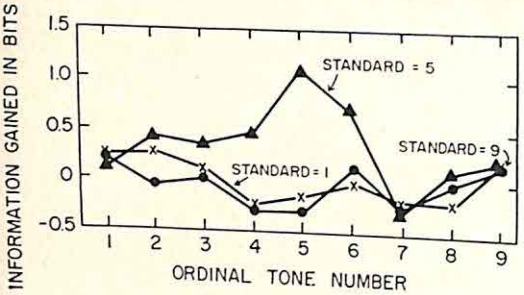


FIG. 5. Gain in information transmitted with objective standard equal in frequency to first, fifth, or ninth tone of the series. (Reprinted with permission from an article by I. Pollack published in the *Journal of the Acoustical Society of America,* 1953, Vol 25. Copyrighted by the Acoustical Society of America, 1953.)

is to provide the subject with objective reference points or standards on each trial. Pollack (1953) used objective reference tones in a study involving absolute judgments of pitch, and his results are shown in Figure 5. Improvement in identification resulting from the presence of the objective standard was found to depend on the position of the standard in the stimulus set. The greatest gain occurred when the standard had a value corresponding to the middle item—that is, when it allowed the subject to partition the stimulus continuum into equal subsets. Moreover, performance was raised not only for the stimulus with the same pitch as the standard, but for adjacent tones as well.

Instead of providing objective standards on each trial, as in Pollack's study, one might attempt to develop mnemonic anchor points, or subjective standards, through training. Cuddy (1968) has demonstrated that training procedures which permit subjects to organize the stimulus set are much more effective than normal absolute judgment practice. Two groups of subjects were given eight training sessions in which they judged 10 tones of varying frequency, and note names were used as responses. One group was given typical absolute judgment practice with feedback for all stimuli. The second group was given feedback for only one item, the musical note "A," which was the middle frequency. For the A-training group, the proportion of A's was progressively reduced from an initial level of .40 as the subjects became more proficient.

The results of Cuddy's experiment are shown in Figure 6. On a pretest, both groups did equally well. But on a posttraining test, the group given special A training greatly surpassed the control group in the identification of all tones, particularly A. It appears that the note "A" acted as an acquired standard which allowed the subjects to bisect the stimulus set into two segments. Instead of showing the normal U-shaped ordinal-position curve, the A training group demonstrated a double-U-shaped function anchored on A. This chunking of the pitch continuum was obviously beneficial, as it raised performance to a much higher level than that found typically in absolute judgment.

Instead of attempting to train subjective standards, as in Cuddy's experiment, one might take advantage of the subject's prior learning history by using specific stimuli which have well-established associations. We have previously characterized such stimuli as high in meaningfulness, and brief mention was made of "bumpy" ordinal position effects which are found when high-

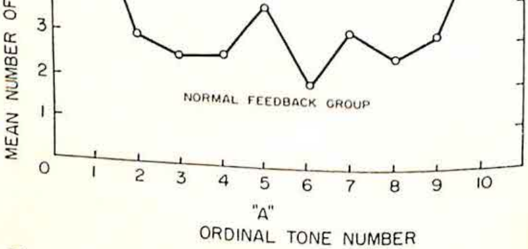


FIG. 6. Mean number of correct judgments per tone in posttraining tests for subjects receiving A training and regular absolute judgment practice. (Reprinted with permission from an article by L. L. Cuddy published in the *Journal of the Acoustical Society of America,* 1968, Vol. 43. Copyrighted by the Acoustical Society of America, 1968.)

meaningful stimuli were used in absolute judgment. An experiment by Weintraub (1971) provides an excellent illustration of how the use of meaningful stimuli may lead to organization of the stimulus set. In a standard absolute judgment task, 15 rectangles varying only in height were used as stimuli, with numbers as responses. Normally, such a set of stimuli would produce the usual bow-shaped ordinal position curve "anchored" at the end points. In this experiment, one of the rectangles in each set was unique in that its height and width were equal. That rectangle could be labeled reliably and uniquely as a *square*. The experimenter did not bring the existence of the square to the subject's attention, and the square was assigned a number response in the same way as the other rectangles in the series, according to ordinal position.

Weintraub's data are shown in Figure 7. It is clear that the square tended to raise the ordinal position curve, and the greatest improvement came when the square was the middle item. Under these conditions, the subject could use the square as a reference point for grouping the stimuli into two subsets, short rectangles and tall rectangles. Weintraub obtained a double-U-shaped ordinal position curve under conditions which facilitated organization, and this is similar to the results of Cuddy's experiment where the stimulus set was bisected by an implicit standard.

We have interpreted the findings of Pollack (1953), Cuddy (1968), and Weintraub (1971) as demonstrating organization or chunking in absolute judgment. An experiment of Spitz (1967) lends considerable support to this view and suggests that organization of a stimulus continuum represents a cognitive strategy. Spitz had groups of retardates and controls judge the position of a pointer on a line in a standard absolute judgment task. As one might suspect, the control subjects performed at a considerably higher level than the retardates, and Spitz attributed this finding to the fact that retardates are less likely to organize input. While the ordinal position curves of the retardates were U shaped, the control group equated for chronological age began to develop a pronounced peak in the middle after about 30 practice trials. It appeared that this group was mentally bisecting the line and forming a conceptual anchor at the midpoint. A second control group equated for mental age with the retardates developed an implicit standard only after being prompted with the presence of an objective standard corresponding to the middle of the line. The retardates showed the expected double-U-shaped curve while an objective standard was present, but failed to maintain an implicit standard after it was removed. Thus, Spitz has demonstrated that the degree of organization of the stimulus continuum in absolute judgment depends on the subject's intellect, and that even with more intelligent observers, practice or familiarity with the stimulus set is required before double-U-shaped ordinal position curves are observed. This is what one would expect if such curves result from a cognitive strategy rather than from some perceptual

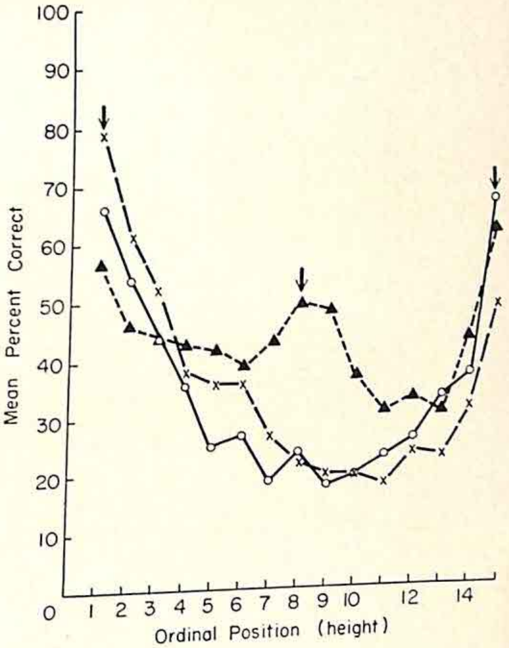


FIG. 7. Mean percent correct in judgment of the height of rectangles (width constant). (In three conditions, either the first, eighth, or fifteenth item was a square.) (Reprinted from an article by D. J. Weintraub published in the *Journal of Experimental Psychology*, 1971, Vol. 88. Copyrighted by the American Psychological Association, 1971.)

factor inherent in the stimulus. It appears then that, given the opportunity, subjects will exhibit the same kinds of organizational behavior in dealing with the sensory stimuli commonly used in absolute judgment as when they are asked to learn verbal materials.

PRACTICE EFFECTS AND PERFECT PITCH

Needless to say, although performance in paired-associate learning is initially quite limited, subjects very quickly learn to identify large numbers of items. On the other hand, the concept of a fixed channel capacity implies that practice effects are minimal in absolute judgment. This myth has persisted despite the very adequate documentation of the role of learning in judgment that exists in the literature (cf. Gibson, 1953; Lawrence, 1963; Neu, 1947; Postman, 1963). One reason for this misconception may be that only a very small number of investigators who have used the absolute judgment procedure have plotted learning effects over trials. When this is done, there is no doubt that the rate of learning is painfully slow (e.g., Alluisi & Sidorsky, 1958; Hartman, 1954), but learning takes place nonetheless.

We would like to think that considerable learning may take place with sensory stimuli, and it would help our argument if we could point to individuals who can identify large numbers of unidimensional stimuli. A striking example comes to mind in those persons who possess "perfect pitch"—the ability to make absolute identifications of musical notes. In talking about the channel capacity, Miller (1956) noted that he was troubled by the existence of such individuals, but decided to ignore the phenomenon of perfect pitch, since he could not explain it. Although perfect pitch seems to contradict the notion of a fixed upper limit on absolute judgment performance, it fits in well with the position presented here.

Perfect pitch has been viewed by several writers (Bachem, 1937; Seashore, 1940) as an inborn gift, inexplicable in terms of learning principles. This position is not in accord with a variety of data which indicate that perfect pitch depends on specific musical experiences (e.g., Boggs, 1907; Stumpf, 1883; Van Krevelen, 1951). Nor can it account for the finding that individuals with perfect pitch are no better at discriminating or remembering pitch differences than control subjects when nonmusical tone frequencies are used as stimuli (Bachem, 1954; J. A. Siegel, 1972).

Until recently, however, those writers who feel that the phenomenon is an innate gift have been able to point to the fact tha experimental attempts to train perfect pitc have not been successful. Now there i evidence that the traditional training procedures are not the most effective ones for low-meaningful sensory stimuli. Cuddy (1968) found substantial improvement in the identification of pitch using a novel training procedure which led to chunking of the pitch continuum, a procedure which we described earlier. Brady (1970) has employed Cuddy's method to acquire perfect pitch. Over a period of a year, Brady listened to and tried to identify randomly presented piano notes, and at the end of this time he was correct to within a semitone on 97% of the trials. This is especially impressive when one considers the conditions of testing; notes were presented one per day early in the morning, and thus the high level of performance that Brady managed to achieve could not have been due to good relative pitch.

It seems reasonable to conclude, then, that the identification of unidimensional stimuli can be greatly improved with training, especially when training leads to organization of the stimulus set, and that individuals with perfect pitch have approached the upper end of the learning curve. Indeed Attneave (1959) points out that subjects with good relative pitch are capable of transmitting three to four bits of information. Thus the level of performance which is usually thought to be typical of absolute judgment—the "magical number seven" phenomenon—in fact represents performance at an early stage of practice. The concept of a fixed upper limit on absolute judgment accuracy appears to be incorrect.

CONCLUDING STATEMENT

In this paper we attempted to demonstrate that absolute judgment and paired-associate learning are highly similar paradigms. Absolute judgment may be characterized as a special case of paired-associate learning in which the stimuli and responses are ordered, and responses are assigned to stimuli in serial order. Not only is there a formal similarity between absolute judgment and paired-associate learning, but they also share a number of common phenomena. Such variables as stimulus meaningfulness, the retention interval, ordinal position, and practice appear to have similar effects in both tasks, despite the fact that the same effect may be given different names and theoretical interpretations in absolute judgment and paired-associate learning.

One result of our attempt at integration is an increased emphasis on learning and memory processes in absolute judgment. Indeed, on the basis of the available data (W. Siegel, 1972), it appears that virtually all of the variance in the typical absolute judgment study is attributable to forgetting. This is somewhat surprising since the absolute judgment task developed within psychophysics, and absolute judgment data have often been used in the development of "sensory" scales. Presumably, psychophysical scales should reflect the rules by which the sensory systems transform physical energy, and not the particular methodology used to obtain them. If this goal is to be realized, memory effects must be unconfounded from phenomena more directly related to the operation of the sensory systems.

Some years ago, Miller (1956) wrote of a magical number seven which seemed to appear quite unexpectedly in a variety of apparently unrelated sets of experimental data. Absolute judgments, rating scales, the subitizing phenomenon, and the span of immediate memory all seemed to be restricted by some common process which had a limited capacity of about seven items. But a careful survey of the literature led Miller to conclude that the span of absolute judgments and the immediate memory span were unrelated. For, he argued, absolute judgment performance is limited by the amount of *information* in the stimuli, while memory span is limited by the number of *items*, and is independent of the information per item. That is, the span of immediate memory does not depend on the size of the set from which the items (e.g., binary digits, decimal digits, letters, or words) are drawn. Moreover, performance within the memory span task could be greatly improved, according to Miller, if subjects organize or chunk the items together. On the other hand, the absolute judgment channel capacity was viewed as "a clear and definite limit to the accuracy with which we can identify absolutely the magnitude of a unidimensional variable [Miller, 1956, p. 90]," and it was thought to be relatively uninfluenced by practice.

With the advantage of hindsight, we can now suggest that Miller's distinction was incorrect. The data of W. Siegel (Figure 3) demonstrate that absolute judgment accuracy is *not* limited by stimulus information, but rather by subjects' inability to maintain multiple representations of sensory stimuli in memory. Furthermore, the hypothesis of a fixed upper limit on absolute judgment accuracy is contradicted by data which demonstrate that identification of unidimensional stimuli may be improved dramatically, provided that the subjects are induced to chunk the stimulus set. Training procedures which emphasize organizational strategies rather than rote learning appear to be highly effective in improving memory for both sensory and verbal stimuli. Thus the absolute judgment channel capacity and the immediate memory span for verbal material both appear to be a limitation on memory during the initial stages of practice, and both may be overcome by organization or chunking. What then of the magical number seven (plus or minus two)? At least in the case of the immediate memory span and the span of absolute judgments, there may be something more involved here than just a "pernicious Pythagorean coincidence."

REFERENCES

Allan, L. G. Visual position discrimination: A model relating temporal and spatial factors. *Perception and Psychophysics,* 1968, **4**, 267–278.

Alluisi, E. A., & Sidorsky, R. C. The empirical validity of equal discriminability scaling. *Journal of Experimental Psychology,* 1958, **55**, 86–95.

Archer, E. J. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs,* 1960, **74**(10, Whole No. 497).

Atkinson, R. C., & Shiffrin, R. N. Human memory: A proposed system and its control processes. In K. W. Spence & Janet T. Spence (Eds.), *The psychology of learning and motivation.* New York: Academic Press, 1968.

Attneave, F. *Application of information theory to psychology.* New York: Holt, 1959.

Bachem, A. Various types of absolute pitch. *Journal of the Acoustical Society of America,* 1937, **9**, 146–151.

Bachem, A. Time factors and absolute pitch determination. *Journal of the Acoustical Society of America,* 1954, **26**, 751–753.

Battig, W. F. Paired-associate learning. In T. R. Dixon & D. Horton (Eds.), *Verbal behavior and general behavior theory.* Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Beare, A. C. Color name as a function of wavelength. *American Journal of Psychology,* 1963, **76**, 248–256.

Boggs, L. P. Studies in absolute pitch. *American Journal of Psychology,* 1907, **18**, 194–205.

Brady, P. T. Fixed-scale mechanism of absolute pitch. *Journal of the Acoustical Society of America,* 1970, **40**, 883–887.

Brelsford, J. W., Keller, L., Shiffrin, R. M., & Atkinson, R. C. Short-term recall of paired-associates as a function of the number of interpolated pairs. *Psychonomic Science,* 1966, **4**, 73–74.

Brown, S. C., & Battig, W. F. Second-list paired-associate facilitation produced by addition of previously learned first-list pairs. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5**, 320–321.

Bush, R. R., Galanter, E., & Luce, R. D. Characterization and classification of choice experiments. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963.

Calfee, R. C. Interpresentation effects in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7**, 1030–1036.

Chapanis, A., & Halsey, R. Absolute judgments of spectral colors. *Journal of Psychology,* 1956, **42**, 99–103.

Conover, D. W. *The amount of information in the absolute judgment of Munsell hues.* (USAF WADC Tech. Note No. 58-262) Washington, D. C.: Government Printing Office, 1959.

Corcoran, E. W. J., Carpenter, A., Webster, J. C., & Woodhead, M. M. Comparisons of training techniques for complex sound identification. *Journal of the Acoustical Society of America,* 1968, **44**, 157–167.

Cuddy, L. L. Practice effects in absolute judgments of pitch. *Journal of the Acoustical Society of America,* 1968, **43**, 1069–1076.

DeSoto, C. B., & Bosley, J. J. The cognitive structure of a social structure. *Journal of Abnormal and Social Psychology,* 1962, **64**, 303–307.

Ebenholtz, S. M. Serial position effect of ordered stimulus dimensions in paired-associate learning. *Journal of Experimental Psychology,* 1966, **71**, 132–137.

Eriksen, C. W., & Hake, H. W. Multidimensional stimulus differences and accuracy of discrimination. *Journal of Experimental Psychology,* 1955, **50**, 153–160.

Eriksen, C. W., & Hake, H. W. Anchor effects in absolute judgments. *Journal of Experimental Psychology,* 1957, **53**, 132–138.

Feigenbaum, E. A., & Simon, H. A. A theory of the serial position effect. *British Journal of Psychology,* 1962, **53**, 307–320.

Garner, W. R. An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology,* 1953, **46**, 373–380.

Garner, W. R. *Uncertainty and structure as psychological concepts.* New York: Wiley, 1962.

Garner, W. R., & Hake, H. W. The amount of information in absolute judgments. *Psychological Review,* 1951, **58**, 446–459.

Gibson, E. J. Improvement in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin,* 1953, **50**, 401–431.

Glaze, J. A. The association value of nonsense syllables. *Journal of Genetic Psychology,* 1928, **35**, 255–267.

Hartman, E. B. The influence of practice and pitch-distance between tones on the absolute identification of pitch. *American Journal of Psychology,* 1954, **67**, 1–14.

Hovland, C. I. Experimental studies in rote-learning theory: III. Distribution of practice with varying speed of syllable presentation. *Journal of Experimental Psychology,* 1938, **23**, 172–190.

Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *American Journal of Psychology,* 1933, **45**, 730–734.

Jensen, A. R. Temporal and spatial effects of serial position. *American Journal of Psychology,* 1962, **75**, 390–400.

Johnson, P. J., Chatelier, P., & Mocharnuk, J. Factor influencing short-term kinesthetic recall. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1969.

Kinchla, R. A., & Allen, L. G. Visual movement perception: A comparison of sensitivity to vertical and horizontal movement. *Perception and Psychophysics,* 1970, **8**, 399–405.

Kinchla, R. A., & Smyzer, F. A diffusion model of perceptual memory. *Perception and Psychophysics*, 1967, **2**, 219–229.

Klemmer, E. T., & Frick, F. C. Assimilation of information from dot and matrix patterns. *Journal of Experimental Psychology*, 1953, **45**, 15–19.

Lawrence, D. H. The nature of a stimulus: Some relationships between learning and perception. In S. Koch (Ed), *Psychology: A study of a science*. Vol. 5. New York: McGraw-Hill, 1963.

Lenneberg, E. H. A probabilistic approach to language learning. *Behavioral Science*, 1957, **2**, 1–13.

Lenneberg, E. H. Color naming, color recognition, color discrimination: A re-appraisal. *Perceptual and Motor Skills*, 1961, **12**, 375–382.

Levett, C. M., Jr. Errors of interpolation in instrument reading and setting. *Journal of Applied Psychology*, 1952, **36**, 49–52.

Mandler, G. Verbal learning. In G. Mandler & P. Mussen (Eds.), *New directions in psychology III*. New York: Holt, Rinehart & Winston, 1967.

Martin, E. Relation between stimulus recognition and paired-associate learning. *Journal of Experimental Psychology*, 1967, **74**, 500–505. (a)

Martin, E. Stimulus recognition in aural paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 272–276. (b)

Martin, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, 1968, **75**, 421–441.

Martin, E., & Melton, A. W., Meaningfulness and trigram recognition. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 126–135.

Massaro, D. W. Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1970, **83**, 32–39.

McGuire, W. J. A multiprocess model for paired-associate learning. *Journal of Experimental Psychology*, 1961, **63**, 81–99.

Miller, G. A. The magical number seven, plus or minus two. *Psychological Review*, 1956, **63**, 81–97.

Muller, P. F., Jr., Sidorsky, R. C., Slivinske, A. J., Allusi, E. A., & Fitts, P. M. *The symbolic coding of information on cathode ray tubes and similar displays*. (USAF WADC Tech. Report No. 55–375) Washington, D. C.: Government Printing Office, 1955.

Murdock, B. B. The distinctiveness of stimuli. *Psychological Review*, 1960, **67**, 16–31.

Nash, H. The judgment of linear proportions. *American Journal of Psychology*, 1964, **77**, 480–484.

Neu, D. M. A critical review of the literature of "absolute pitch." *Psychological Bulletin*, 1947, **44**, 249–266.

Noble, C. E. An analysis of meaning. *Psychological Review*, 1952, **59**, 421–430.

Noble, C. E., Stockwell, F. E., & Pryor, M. W. Meaningfulness (m′) and association value (a) in paired-associate syllable learning. *Psychological Reports*, 1957, **3**, 441–452.

Paivio, A. U. *Imagery and verbal processes*. New York: Holt, Rinehart & Winston, 1971.

Peterson, L. R., Saltzman, D., Hillner, K., & Land, V. Recency and frequency in paired-associate learning. *Journal of Experimental Psychology*, 1962, **63**, 396–403.

Pollack, I. The information of elementary auditory displays. *Journal of the Acoustical Society of America*, 1952, **24**, 745–749.

Pollack, I. The information of elementary auditory displays: II. *Journal of the Acoustical Society of America*, 1953, **25**, 765–769.

Pollack, I., & Ficks, L. Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America*, 1954, **26**, 155–158.

Pollio, H. R. Associative structure and verbal behavior. In T. R. Dixon & D. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice Hall, 1968.

Pollio, H. R., & Draper, D. O. The effect of a serial cognitive structure on paired-associates learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 301–308.

Postman, L. Perception and learning. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 5. New York: McGraw-Hill, 1963.

Prytulak, L. S. Natural language mediation. *Cognitive Psychology*, 1971, **2**, 1–56.

Rubin, E. D., & Brown, S. C. Constant versus varied serial order in paired-associate learning: The effect of formal intralist similarity. *Journal of Experimental Psychology*, 1967, **75**, 257–262.

Seashore, C. E. Acquired or absolute pitch. *Music Educators Journal*, 1940, **26**, 18.

Segal, M. A., & Mandler, G. Directionality and organizational processes in paired-associate learning. *Journal of Experimental Psychology*, 1967, **74**, 305–311.

Siegel, J. A. The nature of absolute pitch. *Studies in the Psychology of Music*, 1972, in press.

Siegel, W. Memory effects in the method of absolute judgment. *Journal of Experimental Psychology*, 1972, **94**, 121–131.

Spitz, H. H. Information transmission in an absolute judgment task with feedback using normal and retarded subjects. *Journal of Comparative and Physiological Psychology*, 1967, **64**, 85–92.

Stumpf, C. *Tonpsychologie*. Leipzig: Hirzel, 1883.

Tanner, W. P., Jr. Physiological implications of psychological data. *Annals of the New York Academy of Science*, 1961, **89**, 752–765.

Thurstone, L. L. Psychophysical analysis. *American Journal of Psychology*, 1927, **38**, 368–389.

Torgerson, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.

Van Krevelen, A. The ability to make absolute judgments of pitch. *Journal of Experimental Psychology*, 1951, **42**, 207–215.

VOLKMAN, F. C., & ENGEN, T. Three types of anchoring effects in the absolute judgment of hue. *Journal of Experimental Psychology*, 1961, **61**, 1–17.

WEBSTER, J. C., CARPENTER, A., & WOODHEAD, M. M. Identifying meaningless tone complexes. *Journal of the Acoustical Society of America*, 1968, **44**, 606–609.

WEINTRAUB, D. J. Rectangle discriminability: Perceptual relativity and the law of prägnanz. *Journal of Experimental Psychology*, 1971, **88**, 1–11.

WEVER, E. G., & ZENER, K. E. The method of absolute judgment in psychophysics. *Psychological Review*, 1928, **35**, 466–493.

WICKELGREN, W. A. Consolidation and retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1966, **72**, 250–259.

WICKELGREN, W. A. Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, 1969, **6**, 13–61.

WITMER, L. R. The association value of three-place consonant syllables. *Journal of Genetic Psychology*, 1935, **47**, 337–360.

YOUNG, R. K. Serial learning. In T. R. Dixon & D. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

YOUNG, R. K., & BUCK, R. E. Ordinal position effects with a two-dimensional stimulus array. *Journal of Experimental Psychology*, 1968, **77**, 161–163.

(Received September 4, 1971; revision received March 29, 1972)

After July 15, 1972, manuscripts should be submitted to the Editor, George Mandler, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

Psychological Review
1972, Vol. 79, No. 4, 317–328

SELECTIVE ATTENTION AND THE BREADTH OF LEARNING:

AN EXTENSION OF THE ONE-LOOK MODEL[1]

BRYAN E. SHEPP,[2] DEBORAH G. KEMLER,[3] AND DANIEL R. ANDERSON[4]

Brown University

When subjects are trained with multiple cues relevant to solution of a discriminative learning or concept identification task, some subjects learn about more than one relevant cue, a result that has been considered damaging to "one-look" versions of attention theory. It is demonstrated that, to the contrary, a slightly extended version of the D. Zeaman and B. J. House one-look model can predict multiple-cue learning in the redundant relevant cues experiment and can accommodate the effects of various task and training factors that have been shown to affect the breadth of learning. Moreover, a new version of the one-look model that assumes all-or-none learning predicts with quantitative rigor the data of T. Trabasso and G. H. Bower on the breadth of learning by adult humans. Thus, the validity of the one-look assumption is by no means undermined by reports of multiple-cue learning in multitrial learning experiments.

The concept of selective attention has been revived by several recent theories of choice behavior to account for discriminative control in a variety of experimental situations (e.g., Lovejoy, 1968; Sutherland & Mackintosh, 1971; Zeaman & House, 1963). Although there are important differences between these theories, they commonly assume that subjects, in solving a discriminative or concept-identification task, may be controlled by or learn about only selected aspects of the total stimulus situation. The results of many different types of experiments have been interpreted as supporting attention theory (e.g., Mackintosh, 1965a; Wolff, 1967), and several major phenomena including blocking (Mackintosh, 1965b), the overtraining reversal effect (Lovejoy, 1966), dimensional learning (Shepp & Eimas, 1964), and progressive improvement in serial reversal learning (Mackintosh, 1969; Shepp & Turrisi, 1969a) have been successfully predicted by the theory.

More recently, several investigators have turned to a consideration of learning in the relevant redundant cues task as a critical test of attention theory. In relevant redundant cues task, more than one dimension is relevant to solution. Consider, for example, a simultaneous discrimination in which a red square (+) and a green circle (−) are presented. The positions (left or right) of the stimuli are irrelevant; red and square are always paired, as are green and circle. After the task has been learned, a series of test trials or a transfer task is given to provide a basis for inferring what has been learned about each dimension.

The most intuitive prediction of attention theory about learning in a relevant redundant cues task is that the more a subject learns about the cues on one relevant dimension, the less he learns on the second. This prediction follows most directly from "one-look" models of attention theory (e.g., Lovejoy, 1968; Zeaman & House, 1963) which assume

---

[1] Preparation of this paper was supported in part by Research Grants HD 04320 and MH 16642 from the United States Public Health Service awarded to the first author, by a National Science Foundation Predoctoral Fellowship awarded to the second author, and by Office of Education Grant OEG-1-71-0108 508 awarded to the third author. The authors wish to express their appreciation to Maro Dermateosian and Shellie Hoffer for their assistance in running the simulations.

[2] Requests for reprints should be sent to Bryan E. Shepp, Walter S. Hunter Laboratory of Psychology, Brown University, Providence, Rhode Island 02912.

[3] Now at the University of Pennsylvania, Department of Psychology, Philadelphia, Pennsylvania 19104.

[4] Now at the University of Massachusetts, Bartlett Hall, Amherst, Massachusetts 01003.

that attention and learning are limited to a single dimension on any single trial, and that, with sufficient training, attention is consistently controlled by the same dimension on every trial.

This prediction has been confirmed in a series of experiments by Sutherland and Holgate (1966). Rats were trained on a relevant redundant cues task, and then were administered a series of test trials in which only one of the relevant dimensions could serve as a basis for choice. In most of their experiments, Sutherland and Holgate report a negative correlation between the number of responses to the correct cue on one dimension and the number of responses to the correct cue on the other. According to a one-look interpretation, the negative correlation is due to the increasing restriction of attention to only one dimension during the course of training. The restriction of attention to one relevant dimension allows learning to continue on that dimension, but minimizes additional learning on a second relevant dimension.

There seems to be little doubt that a one-look model can predict a negative correlation between the amount learned about one set of relevant cues and the amount learned about a second set. Some experiments, however, have failed to find a negative correlation and have reported instead considerable learning on both dimensions by individual subjects, a result that some (e.g., Honig, 1969; Stollnitz, 1969; Warren & McGonigle, 1969) have interpreted as inconsistent with a one-look model. In this vein, the most serious criticisms of the one-look interpretation have been made by Trabasso and Bower (1968). They conducted a series of experiments in which human adults were taught concept-identification tasks which had two relevant dimensions. In a series of test trials, some subjects, "one-cue learners," consistently chose the correct cue on one of the relevant dimensions but showed chance performance on the other, whereas other subjects, "two-cue learners," consistently chose the correct cue on both relevant dimensions. The observation of two-cue learning led Trabasso and Bower (1968) to reject their own all-or-none one-look model (Bower & Trabasso, 1964) and to suggest that incremental one-look models (e.g., Lovejoy, 1968; Zeaman & House, 1963), at least in their present forms, were inadequate to predict the observation of two-cue learning. In place of such models, Trabasso and Bower (1968) have proposed a "multiple-look" model which assumes that while attention is selective (i.e., not *all* cues are attended to on a single trial), it is broader than is assumed by one-look views.

In our opinion, the phenomena of one- and two-cue learning in human experiments and the analogous correlational results in the animal literature are entirely consistent with a one-look interpretation. In the present paper, we support this view by providing an analysis of attention and learning in relevant redundant cue situations according to a slightly extended version of Zeaman and House (1963), and we show that this model can accommodate the major results of relevant redundant cue experiments. Further, we propose a novel version of an all-or-none one-look model that can predict the quantitative details of Trabasso and Bower's (1968) results as accurately as does their own multiple-look model.

Zeaman-House One-Look Model

According to Zeaman and House (1963), the subject is presented with a stimulus situation that can be analyzed into k relevant and irrelevant dimensions. On each trial, exactly one of the k dimensions elicits attention just prior to choice. Attention to Dimension i occurs with probability d_i, and, in a simultaneous discriminative task, exposes the cues $C_{i,1}$ and $C_{i,2}$. The subject may choose either cue with the conditional probabilities $c_{i,1}$ and $c_{i,2}$, where $c_{i,2} = 1 - c_{i,1}$. The probabilities, d_i and $c_{i,1}$, are modified by trial outcomes. Reinforcement following attention to Dimension i increases d_i, whereas nonreinforcement decreases d_i. Moreover, on all trials, the k attentional probabilities must sum to 1.0, so that reinforcement or nonreinforcement of attention to one dimension, respectively, decreases or increases the probabilities of attention to other dimensions. Similarly, reinforcement following the choice of a cue increases the conditional probability of choosing that cue, and nonreinforcement decreases that probability. A change in $c_{i,1}$ leads to a complementary change in $c_{i,2}$; correspondingly, a change in $c_{i,2}$ produces a complementary change in $c_{i,1}$.

A critical feature of the one-look model is that reinforcement or nonreinforcement on any trial affects only the probabilities of choosing the specific cues on the dimension to which the subject attends on that trial; the probabilities of choosing cues on dimensions that are ignored do not change. For example, if a subject attends to color on a given trial, the probabilities of choosing the specific color cues will be modified, but the conditional probabilities of choosing cues on other dimensions (e.g., right or left position) will not be changed on that trial.

Consider now the implications of these assumptions in the learning of the relevant redundant cues task described previously. A subject might attend only to form in a series of training trials, and in so doing learn only that square is correct and circle is incorrect. This subject would learn nothing about the correctness of red and green, and would be characterized as a one-cue learner. The theory (Zeaman & House, 1963) does not demand, however, that the subject attends to the *same* dimension on *every* trial. On some trials, the subject may attend to form and learn about the form cues; on other trials, the subject may switch attention to color and learn the correct response assignments to red and green. In this event, the subject has clearly learned about the correct cues on both dimensions. Yet, for both one- and two-cue learners, the theory predicts consistent control by one dimension on conflict trials in which the positive cue on each relevant dimension is paired with the negative cue on the other; according to a one-look model, the probability of attending to one relevant dimension approaches unity as training progresses, so that subjects will respond to only one of the relevant cues in a series of conflict trials introduced after a strong criterion run. Such performance is typically observed (e.g., Cole, 1953; Trabasso & Bower, 1968).

While the conflict type of test trial is useful in assessing which relevant dimension controls the discriminative response, a so-called component type of test trial is necessary to assess what has been learned about each of the relevant dimensions. In the component test trial, one relevant dimension is neutralized or eliminated, so that performance on the other can be observed. The Zeaman-House model is silent about performance on component test trials, since no rules have been provided for specifying the distribution of attention probabilities when a dimension of variation is eliminated from the stimulus situation.[5] A small, but crucial, extension is required. Following the suggestion of Trabasso and Bower (1968), the deficiency has been amended by adopting Luce's (1959) choice axiom. When Dimension i is eliminated, the probabilities of attending to all other dimensions are renormalized, such that the probability of attending to D_j on a test trial when D_i is removed, $d_{j(\bar{i})}$, may be specified as

$$d_{j(\bar{i})} = \frac{d_j}{1 - d_i},$$

where d_j and d_i refer to the probabilities of attending to Dimensions i and j after the preceding training trial. Given this assumption, the model can predict two-cue learning despite the assumption that attention is entrapped by only one of the dimensions in a series of training trials.

Consider again the subject who learns the correct response assignments to both color and form over a series of relevant redundant cues training trials. Given sufficient training, such a subject will come exclusively under the control of either the color or the form dimension. The combined probability of attending to the other relevant dimension and all irrelevant dimensions (e.g., position in the example of a simultaneous discrimination) may well be less than .01 after an extended criterion run. Despite this fact, the probability of attending to the nondominant relevant dimension may remain relatively high in comparison with the probability of attending to the irrelevant dimensions, provided that some learning has occurred on the nondominant dimension. Consequently, when the dominant dimension is removed from a test trial, the probability of attending to the nondominant dimension may be very high such that whatever learning has occurred on this dimension will be reflected in the test trial performance. Thus, contrary to the claim of Trabasso and Bower (1968) and others, two-cue learning is by no means a rare prediction of a one-look model. The principle issue now becomes how well this modified version of the Zeaman-House model can predict the results of relevant redundant cues experiments.

---

[5] It should be noted that Lovejoy's (1968) model suffers the same deficiency. He has specified no rules for determining the distribution of "directable distinctiveness," given the addition or deletion of a dimension.

Relevant Redundant Cues Learning

The literature on relevant redundant cues learning is not extensive, but several principles seem to be emerging. First, the observed proportions of one- and two-cue learners are affected by the relative saliences of the relevant cues, with the proportion of two-cue learners increasing as the saliences of the relevant cues converge (Trabasso & Bower, 1968). Second, the breadth of learning is increased by certain conditions of practice. The administration of overtraining (Sutherland & Holgate, 1966) or partial reinforcement (Sutherland, 1966) increases the amount of two-cue learning.

In order to demonstrate that the Zeaman-House model can accommodate the results of relevant redundant cues studies, it is necessary to simulate the performances of stat subjects under the various conditions of interest. In the simulations that follow, each condition contains 15 subjects. Each subject is run to a learning criterion of nine correct responses in 10 trials. Unless it is otherwise specified, three dimensions are assumed to be effective; two are relevant, one is irrelevant. The rate parameters for attentional responses (θ_o) and instrumental responses (θ_r) are set at .2 and .5, respectively. Although these values were chosen arbitrarily, they fall well within the range of values used previously to fit experimental data (e.g., Shepp & Turrisi, 1969b; Shepp & Zeaman, 1966; Zeaman & House, 1963). The probability of choosing either cue of a dimension is set initially at .5 for all dimensions so as to correspond to a two-choice situation. In all simulations, a two-cue learner is defined as a stat subject who is expected to make at least seven correct responses in each series of 10 test trials.[6]

Salience of cues. There are many operations by which one set of cues can be made more salient than another set, and some of these have been described by Trabasso and Bower (1968). By varying the salience of cues, one may vary the initial probabilities of attending to specific dimensions, with more salient cues having a higher probability of being attended to than less salient cues (e.g., Shepp & Zeaman, 1966; Trabasso & Bower, 1968). In the relevant redundant cues situation, the proportions of one- and two-cue learners can vary dramatically with the relative salience of the redundant relevant dimensions.

The first illustration of this effect is shown in Table 1. In this case the total salience of the relevant dimensions is held constant in three different conditions, with the initial probabilities of attending to the relevant dimensions summing to .90. Under this condition the amount of two-cue learning increases as the saliences of the relevant dimensions converge ($\chi^2 = 6.96$, $df = 2$, $p < .05$). The amount of two-cue learning also increases as the total salience of the relevant dimensions increases. This relationship is shown in Table 2. The initial probabilities of attending to the relevant dimensions are set at a constant ratio of 2:1, and the sum of these probabilities is varied from

---

[6] In generating specific predictions for test trial performance, it is assumed that attentional and instrumental probabilities remain stable over the test trial series. This is only the ideal case, but there is empirical evidence that this stability can be approximated at least with human subjects (Eimas, 1969; Trabasso & Bower, 1968). Although the predicted proportions of one- and two-cue learners would be changed by assuming that nondifferential reinforcement or nonreinforcement is effective on the test trials, the relative predictions for the various conditions of interest would not be modified and the conclusions would be the same.

.075 to .90. It is clear from an inspection of Table 2 that the amount of two-cue learning increases with total salience of the relevant dimensions. This conclusion is confirmed statistically ($\chi^2 = 8.23$, $df = 2$, $p < .02$) and supported empirically by Trabasso and Bower (1968) and by Miles and Jenkins (1965).

The number of dimensions. Generally speaking, the addition of dimensions to a revelant redundant cues task results in less being learned about any one relevant cue regardless of whether the added dimensions are relevant or irrelevant. Exceptions to this conclusion may occur if one of the dimensions is much more salient than the others. Under such conditions, however, the number of dimensions makes little sense as an independent variable. To illustrate the effect of adding dimensions to a relevant redundant cues task, the base-level comparison is the amount of one- and two-cue learning in tasks in which there are but two equally salient relevant dimensions.

Consider the effect of adding either one or three irrelevant dimensions to the task. All dimensions are assumed to be equally salient: In the three-dimensional problem, the initial probabilities of attention are .33 for each dimension; in the five-dimensional problem, the probabilities are .20. By adding irrelevant dimensions, the total salience of the relevant dimensions is decreased, with the expectation that the number of two-cue learners is similarly reduced.

This expectation is confirmed in Table 3 which shows that significantly less two-cue learning occurs as the number of irrelevant dimensions increases ($\chi^2 = 16.46$, $df = 2$, $p < .001$). In these simulations, the greatest reduction in two-cue learning occurs with the addition of just one irrelevant dimension. While this particular result is surely due, in part, to the choice of parameters, it should be noted that a parallel effect has been observed by Trabasso and Bower (1968, p. 166). In their experiment, the greatest reduction in two-cue learning occurs with the addition of one irrelevant dimension.

TABLE 1

Frequency of One- and Two-Cue Learning with Total Salience of the Relevant Dimensions Held Constant

Initial probability[a]		Type of solution	
D_1	D_2	One cue	Two cue
.85	.05	12	3
.60	.30	7	8
.45	.45	5	10

[a] D_1 and D_2 refer to the relevant dimensions.

TABLE 2

Frequency of One- and Two-Cue Learning with Total Salience of the Relevant Dimensions Varied

Initial probability[a]		Type of solution	
D_1	D_2	One cue	Two cue
.60	.30	7	8
.40	.20	11	4
.05	.025	14	1

[a] D_1 and D_2 refer to the relevant dimensions.

When relevant dimensions are added to the stimulus situation, the total salience of the relevant dimensions increases, but less is learned about the cues on each dimension. If, of course, one of the relevant dimensions is highly salient in comparison with the others, little two-cue or multiple-cue learning will be observed. If, on the other hand, the dimensions are equally salient, the amount learned about any one cue decreases as the number of relevant dimensions is increased. This conclusion agrees with the results of Eimas (1969), and the effect is shown in Table 4 where the average probability of choosing the correct cue in test trials is shown for tasks in which two,

TABLE 3

Frequency of One- and Two-Cue Learning with the Number of Irrelevant Dimensions Varied

Irrelevant dimensions	Type of solution	
	One cue	Two cue
0	2	13
1	11	4
2	12	3

TABLE 4

AVERAGE PROBABILITY OF CHOOSING THE CORRECT CUE ON EACH RELEVANT DIMENSION WITH INCREASING RELEVANT DIMENSIONS

Average probability[a]	Number of relevant dimensions		
	2	3	4
$C_{1,1}$	.930	.932	.758
$C_{1,2}$	.917	.863	.757
$C_{1,3}$		.907	.772
$C_{1,4}$			.832

[a] C_1 designates the correct cue on each dimension.

three, or four dimensions are relevant. Although there is no reduction in the amount learned about Dimensions 1 and 2 in the three-dimensional task, there is a significant reduction in learning in the four-dimensional task ($p < .01$). Similarly, there is less learned on Dimensions 1, 2, and 3 in the four-dimensional than in the three-dimensional task ($p < .01$).

Conditions of practice. Two variables, partial reinforcement and overtraining, have been shown to affect the breadth of learning. Sutherland (1966) compared two groups of rats on a relevant redundant cues task, one receiving the typical differential reinforcement (i.e., 100% vs. 0%) and the other partial reinforcement (i.e., 50% vs. 0%) on the relevant cues. Performances during the test trials indicated that the latter group had learned about more relevant cues than did the former. That this effect can be explained by the switching of attention from trial to trial has been shown by Lovejoy (1968), who has derived these effects for his own one-look model. The predictions from Zeaman and House (1963) are similar and are not considered in this discussion.

Sutherland and Holgate (1966) were the first to show that increases in the amount of training increase the breadth of learning. A group of rats trained to criterion showed less two-cue learning than a group that was given a series of overtraining trials. Trabasso and Bower (1968) have reported a similar effect with the human adult, while Eimas (1969) has found no such effect in children. According to a one-look interpretation (Zeaman & House, 1963), overtraining is likely to increase the breadth of learning if at the beginning of overtraining (*a*) attention to both relevant dimensions is less than asymptotic and (*b*) learning on one of the relevant dimensions is less than asymptotic. If these conditions prevail, the subject may switch attention from trial to trial in overtraining, with the result that additional two-cue learning can occur. Thus, the effect of overtraining on the breadth of learning depends on such factors as the initial salience of the relevant dimensions, the strength of the criterion, and the rate parameters, θ_o and θ_r.

Our initial attempt to simulate the overtraining effect failed due to rate parameters that were too large. With $\theta_o = .2$ and $\theta_r = .5$, the effects were in the right direction, but with these parameters attention is sometimes entrapped at criterion. In other cases too much learning has occurred on both dimensions prior to overtraining such that no effect can be observed. The simulations that follow were generated with $\theta_o = .05$ and $\theta_r = .20$, and with three effective dimensions. It should be noted before proceeding that the previously described relations between salience and the breadth of learning are substantially the same with the lower rate parameters.

The effects of overtraining on learning in the relevant redundant cues task are summarized in Table 5 for several conditions in which the salience of the dimensions is varied.

TABLE 5

FREQUENCY OF ONE- AND TWO-CUE LEARNERS AS A FUNCTION OF TRAINING AND RELATIVE SALIENCE

Condition	D_1	D_2	Training			
			Criterion type of solution		Overtraining type of solution	
			One cue	Two cue	One cue	Two cue
A	.09	.09[a]	8	3	2	13
B	.17	.01	15	0	15	0
C	.45	.45	3	12	1	14
D	.85	.05	12	3	6	9

Note.—D_1 and D_2 are relevant.
[a] Four subjects were "zero" cue learners by our definition.

Criterion subjects were run until nine correct responses in 10 trials occurred; overtrained subjects received an additional 100 trials. An inspection of these results indicates that both the total salience and the ratio of saliences combine with overtraining to affect breadth of learning. When the total salience of the relevant dimensions is relatively low, overtraining increases the amount of two-cue learning, but only as the saliences of the relevant dimensions approach equality (Condition A, $\chi^2 = 7.12$, $df = 1$, $p < .01$). As saliences diverge, only one-cue learning obtains. In contrast, when the total salience is high, overtraining increases two-cue learning but only in the case where the saliences of the relevant dimensions diverge (Condition D, $\chi^2 = 5.00$, $df = 1$, $p < .025$). A convergence of the salient dimensions (Condition C) results in a majority of two-cue learners at criterion. The results of these simulations indicate that overtraining may not always increase the breadth of learning, and suggest that salience may have some role in reconciling those experiments that report an effect due to overtraining with those in which no effect is observed.

It is important to note that overtraining can increase the amount of two-cue learning in a group of subjects only if errors occur during overtraining. In Conditions A and D (Table 5), the percentages of errors for the groups were 7.7 and 3.2, respectively. These error rates in overtraining are in line with the 5% error rate reported by Sutherland and Holgate (1966), who showed that overtraining increases two-cue learning (Experiment 2), and this fit clearly disputes their conclusion that such a low error rate is damaging to a one-look interpretation.

As the modified one-look model can predict that overtraining increases multiple-cue learning in some relevant redundant cues tasks, so also can the model predict incidental learning (e.g., Bitterman & Coate, 1950). According to the one-look model, some incidental learning can occur if the probability of control by the originally relevant dimension is less than asymptotic when the redundant incidental cues are introduced. Once again, the model predicts that within the group, some errors must occur in the redundancy phase in order for incidental learning also to be shown, but the percentage of such errors may be quite small. The one-look model further predicts the reduction of incidental learning by overtraining prior to the addition of the incidentally relevant dimension, an effect that Bruner, Matter, and Papanek (1955) have obtained.

The negative correlation. Sutherland and Holgate (1966) and other investigators of animal behavior rarely present their results in terms of the proportions of one- and two-cue learners. The choice data obtained from animals is almost never unambiguous in this regard, since animals seldom show either perfect performance or exactly chance performance over a series of test trials. Thus, animal investigators have used the correlation between the number of correct responses on each of the component test trials as a test of the one-look model.

Obviously, the more one-cue learning that occurs in relevant redundant cues task, the more highly negative the correlation between test trial performances should be. But just as the one-look model predicts two-cue learning, so too can it predict an attenuation of the correlation of test trial choices. Any condition that favors a high degree of two-cue learning will also attenuate the strength of the negative correlation. For example, in the preceding simulations, overtraining was observed to increase the frequency of two-cue learning (Condition D, Table 5). Correspondingly, at criterion the correlation between test trial performances was $-.75$ ($p < .01$), but after overtraining the correlation shrunk to $-.27$ ($p > .05$).

Although the correlation between choices on test trials can vary considerably and in so doing can reflect the breadth of learning, the strength of the correlation cannot be taken as indicating the degree to which the probability of attending to one relevant dimension is correlated with the probability of attending to a second, as some investigators (e.g., Honig, 1969; Thomas, 1969) have erroneously concluded. By assumption attentional probabilities or analyzer strengths sum to 1.0. Thus, after a series of training trials there must always be a negative correlation between attentional probabilities of relevant dimensions, even when the correlation between choices on test trials is markedly attenuated. To illustrate, consider Condition A of the preceding simulation (see Table 5). After overtraining is administered, the correlation between relevant attentional probabilities is −.99 while the correlation between correct responses on component test trials is −.04.

In addition to those task variables that may affect the breadth of learning during the course of training on a relevant redundant cues task, the type of test trial administered can also affect the breadth of learning. According to the present version of the Zeaman-House model, the probability of attending to a relevant dimension on a test trial depends strongly on the probabilities of attending to other dimensions that are present, and is independent of the probability of attending to the dimension or dimensions that have been eliminated from the test trial. It follows, then, that the number of irrelevant dimensions on test trials can affect the observed proportions of one- and two-cue learning. Trabasso and Bower (1968), for example, retained irrelevant dimensions during test trials and consequently may have minimized observed two-cue learning. In contrast, one could eliminate the irrelevant dimensions from the test trial. This would result in an increase in the probability of attending to the relevant dimension and would result in an attenuation of a negative correlation between test trial performances. Unfortunately, no published data are available to evaluate this implication of the model.

An all-or-none one-look model. In simulating the predictions of the Zeaman-House model, there has been no attempt to predict the numerical details of specific data. Many results can be predicted within a range of parameter values, but it would be a prohibitive task to discover a unique combination of values that would predict all effects. In fact there are $N + r + 1$ free parameters (where N is the number of dimensions and r is the number of relevant dimensions) that need to be specified in order to perform a simulation of the model. If there are 10 "reasonable" values for each parameter, then 10^{N+r+1} simulations would be required to completely explore the parameter space. For example, given two seconds per simulation, more than 63 years of computer time would be required to specify the possible predictions in the situation where $N = 4$ and $r = 4$. Instead, we have qualitatively derived the principal results of relevant redundant cues learning, leaving the problem of detailed quantitative prediction until more adequate procedures for estimating parameter values are developed.

By a few simplifying assumptions, however, we have found it possible to use the one-look model to predict most of the numerical details of those results that Trabasso and Bower (1968) have offered in support of their multiple-look model. By assuming that both attentional and instrumental learning for a given dimension are completed in a single trial, the one-look model becomes mathematically analyzable and allows for specific numerical predictions (Kemler & Anderson, 1971). Although the all-or-none assumption possibly precludes applying the model to data from animal and child subjects, it appears reasonably well fitted to adult human discriminative learning.

This new one-look model may be stated in a very general form. Like Zeaman and House (1963), it assumes that a subject on any trial samples one of the k dimensions of the stimulus situation. The sampling probabilities on Trial 1 are assumed to depend on the relative salience of the dimensions. On Trial n ($n > 1$) the following sampling rules apply: (*a*) if D_i is sampled on Trial $n - 1$, and if the subject responds correctly, then $d_{i,n} = 1$, and $d_{j,n} = 0$, $i \neq j$; (*b*) if D_i is sampled on Trial $n - 1$, and if the subject responds incorrectly, then $d_{i,n} \leq d_{i,n-1}$. In the specific submodel to be presented here, $d_{i,n}$, following an error, is assumed to equal $d_{i,1}$, but other submodels (see Kemler & Anderson, 1971) deal equally well with the results of Trabasso and Bower (1968).

All submodels adopt the following response rules. When any dimension i is first sampled, the probability of choosing $\mathrm{C}_{i,1}$ on that dimension is $c_{i,1}$, and the probability of choosing $\mathrm{C}_{i,2}$ is $1 - c_{i,1}$. However, once a dimension has been sampled, the response to the specific cue becomes conditioned according to the reinforcement outcome of the trial on which the dimension was sampled. Thus, if color is sampled and the choice of red is reinforced, the resampling of color will result in the choice of red with a probability of 1.00. As a consequence of this latter rule, once a relevant dimension is sampled, then regardless of the outcome the conditional probability of choosing the correct cue on a subsequent trial will be 1.00.

To predict test trial performances, the general model adopts Luce's choice axiom and one other assumption as well. On the first test trial that the controlling dimension is removed, the subject resamples from among the remaining dimensions and generates a response on the basis of the selected dimension. On all succeeding test trials of the same type, it is assumed that the subject responds in a manner that is consistent with the choice of the initial trial. Thus, a subject may appear as a two-cue learner by the test trial measures if he sampled the noncontrolling relevant dimension on any earlier trial[7] and resamples it on the test trials, or if he initially samples the noncontrolling relevant dimension on the test trials and guesses the correct response classification along it.

In Chapters 3 and 4 of their book, Trabasso and Bower (1968) make specific predictions from their multiple-look model for a relevant redundant cues experiment and a transfer situation. It is to the results of these experiments that the present one-look model is applied. The crucial prediction for the relevant redundant cues experiment is the proportion of one- and two-cue learners, and in Table 6 the observed proportions along with those predicted from each model are presented. For both models, estimates of the initial saliences of the two relevant dimensions, shape and dot position, are made from the error data of subjects who are trained on a single relevant cue task in which each of the dimensions alone is relevant. These estimates, shown in Table 6, are then used to calculate the proportions of one- and two-cue learners. An inspection of Table 6 reveals that not only does each model predict the results with impressive accuracy, but the models predict equally well ($\chi^2 = 1.72$, $df = 2$, $p > .05$).

TABLE 6

PREDICTED AND OBTAINED PROPORTIONS OF ONE- AND TWO-CUE LEARNERS IN TRABASSO AND BOWER'S (1968) RELEVANT REDUNDANT CUES STUDY

Item	One-look model	Trabasso and Bower model	Obtained data
Salience estimates			
Shape dimension	.135	.094	—
Dot dimension	.218	.164	—
Observed solution types			
Shape alone	27.5	28.8	31
Dot alone	50.3	49.3	43
Both shape and dot	11.2	10.9	15

Consider now Trabasso and Bower's transfer experiment, the results of which are shown in Table 7. The subjects were first trained on a relevant redundant cues task in which shape and dot position were relevant. They were then transferred to a task in which either shape or dot position was retained as relevant, and the other previously relevant dimension was made irrelevant. Data were also obtained from two other transfer conditions which served as controls. One group was first trained with shape as relevant and dot position as irrelevant; in transfer, shape was made irrelevant and dot position became relevant. A second group received the converse arrangement.

In Table 7 the expected numbers of total errors, $E(T)$, that are predicted by each model for the transfer conditions are presented, along with the observed data for each group. In addition, since both the multiple-look and the present one-look models assume that some specific proportion of the subjects,

---

[7] By the assumptions of the model the subject must have made an error on that trial; otherwise he would have gone directly to solution with that dimension controlling his behavior. If subjects tend to be "locally consistent" as Trabasso and Bower (1968) assume, such errors are impossible after Trial 1. However there are data that question such a strict assumption (e.g., Kenoyer, 1972; Levine, 1969) even for adult humans, and the present model allows that these errors do occur. Still, the present model does have the problem of underpredicting local consistency while Trabasso and Bower overpredict it.

TABLE 7

PREDICTIONS AND OBTAINED TRANSFER DATA

Type of transfer task	$E(T)$ Kemler & Anderson	$E(T)$ Trabasso & Bower	Obtained mean total errors	$E(T\|T \geq 1)$ Kemler & Anderson	$E(T\|T \geq 1)$ Trabasso & Bower	Obtained $E(T\|T \geq 1)$
RRC - SRC						
Shape relevant	6.53	7.41	6.62	10.35	11.74	11.31
Dot relevant	2.29	2.66	2.78	6.22	7.18	6.18
Control						
Shape relevant	11.13	11.65[a]	11.74	—	—	—
Dot relevant	6.87	7.10[a]	7.18	—	—	—

Note.—RRC = relevant redundant cues; SRC = single relevant cue.

[a] These estimates are based on a report of $E(T)$ for the single relevant cue groups given in Chapter 4 of Trabasso and Bower (1968) which is inconsistent with the data given in Chapter 3. Employing the data given in Chapter 3, one estimates these values as 11.13 and 6.87, respectively, for Trabasso and Bower's model.

tranferred from relevant redundant cues to single relevant cue tasks, make zero errors, predictions of the expected number of total errors that are conditional on at least one error, $E\ (T|T \geq 1)$, are also shown in Table 7. As may be seen, both the multiple-look and one-look models predict performance in transfer equally well, with the predicted values of each model falling within one standard error of the mean.

Summary and Conclusion

The present analysis of attention and learning shows that a modified version of Zeaman and House (1963) not only can predict one- and two-cue learning, but can also accommodate the major results of the relevant redundant cues literature. One-look models also predict other major phenomena of discriminative learning, including dimensional learning, the overtraining reversal effect, and progressive improvement during serial reversal learning in both animal (e.g., Lovejoy, 1968; Sutherland & Mackintosh, 1971) and human experiments (e.g., Shepp & Turrisi, 1966; Wolff, 1967). Add to these achievements the fact that an all-or-none version of the model predicts the details of human relevant redundant cues data with accuracy, and the resulting picture depicts the one-look model as providing an impressive account of choice behavior. There are, of course, other empirical arguments against a one-look interpretation that have been clearly described by Sutherland and Mackintosh (1971) which we have not considered in the present paper. Instead of attempting to rebut such specific arguments, a task that would require considerable space and several post hoc assumptions, we choose instead to leave the reader with two general points.

First, both one-look and multiple-look models have thus far dealt with the discriminative process in a relatively crude descriptive fashion. As a consequence, it may be impossible to distinguish between the two types of models accurately until some processes are more clearly identified, and attempts to so distinguish them may lead to erroneous conclusions about the correctness of one or the other model. By way of illustration, no attentional model has effectively come to grips with the problems of perceptual organization and the relation of perceptual input to these models. If we are to decide whether a subject has attended to one or more dimensions on a trial, then we must assess independently the number of psychological dimensions that are available to the subject. Two or more experimenter-defined dimensions may be treated as a single pattern by the subject in which case no unambiguous tests of the models are possible. The importance of specifying the effective stimulus has been clearly demonstrated by Garner (1970) in the context of issues raised by information-processing theories.

Unfortunately, the effective psychological dimensions are seldom identified. For example, consider an experiment by Terrace (1963), who showed that pigeons learned about orientation cues while their choices were presumably under the control of color. Sutherland and Mackintosh (1971), among others, have interpreted this finding as contrary to a one-look model, and as requiring some form of multiple-look interpretation. This conclusion is correct, however, only if orientation and color were operating as perceptually independent dimensions, a condition that may not obtain in Terrace's study. The birds showed appreciable learning about orientation cues only under the arrangement where color was gradually faded, which suggests that the birds were not reacting to the two sets of stimuli as perceptually independent dimensions. It is quite plausible that the subjects were attending to cue patterns or compounds that were made salient by the fading procedure.

Although there is no direct evidence to implicate the learning of compound cues in Terrace's situation, there is evidence that animals (e.g., Guth, 1967) and humans (e.g., House & Zeaman, 1963) can learn about such cues. Moreover, there is evidence that while sets of stimuli are perceived as independent dimensions in one type of situation, they are not so perceived in another (e.g., Shepp & Gray, 1971). Thus, the results of Terrace and other results cited by Sutherland and Mackintosh (1971) as contrary to a one-look model may be pointing to a general weakness of all learning models in failing to deal with questions of perceptual inputs and the organization of such inputs. If so, such results cannot distinguish between one-look and multiple-look models.

Our final point is one of theoretical strategy. While it is reasonable to assume that some of the evidence considered as contrary to a one-look model can be eventually accommodated, it is extremely likely that a multiple-look model will also be required. Indeed, it is difficult to suppose that a subject cannot learn to attend to more than a single dimension, especially in cases where either the capacities of organisms are tapped or demands of the task require multiple looks. Yet the frequency of those results demanding a multiple-look interpretation may be relatively small since, as we have shown in the present paper, the one-look model accounts for most of the observed data. Rejection of the one-look model is certainly premature and can, in the long run, lead to erroneous analyses of choice behavior. Strategically it seems wise to retain both models. Both can easily coexist in a general theory of attention, and both may be required to the extent that the theory assimilates the developmental and comparative differences in choice behavior across situations with diverse task demands. However, because we believe that attentional concepts will serve to delineate and clarify important developmental and comparative differences, we urge caution in interpreting choice data. The present paper has shown that data from relevant redundant cues experiments that have appeared to some writers as providing a straightforward index of attentional processes do not on closer analysis fulfill that promise.

REFERENCES

Bitterman, M. E., & Coate, W. B. Some new experiments on the nature of discrimination learning in the rat. *Journal of Comparative and Physiological Psychology*, 1950, **43**, 198–210.

Bower, G. H., & Trabasso, T. Concept identification. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Stanford: Stanford University Press, 1964.

Bruner, J. S., Matter, J., & Papanek, M. L. Breadth of learning as a function of drive level and mechanization. *Psychological Review*, 1955, **62**, 1–10.

Cole, J. The relative importance of color and form in discrimination learning in monkeys. *Journal of Comparative and Physiological Psychology*, 1953, **46**, 16–18.

Eimas, P. D. Multiple-cue discrimination learning in children. *Psychological Record*, 1969, **19**, 417–424.

Garner, W. R. The stimulus in information processing. *American Psychologist*, 1970, **25**, 350–358.

Guth, S. L. Patterning effects with compound stimuli. *Journal of Comparative and Physiological Psychology*, 1967, **63**, 480–485.

Honig, W. K. Attentional factors governing the slope of the generalization gradient. In R. Gilbert and N. S. Sutherland (Eds.), *Animal discrimination learning*. London: Academic Press, 1969.

House, B. J., & Zeaman, D. Miniature experiments in the discrimination learning of retardates. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in child development and behavior.* Vol. 1. New York: Academic Press, 1963.

KEMLER, D. G., & ANDERSON, D. R. *The breadth of attention in learning: A new one-look model.* (Tech. Rep. 5) Providence, R.I.: Brown University Child Study Center, August 1971.

KENOYER, C. E. Experimental tests of models that assume local consistency and resampling only on error trials. *Journal of Experimental Psychology*, 1972, **92**, 215–219.

LEVINE, M. Neo-noncontinuity theory. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 3. New York: Academic Press, 1969.

LOVEJOY, E. Analysis of the overlearning reversal effect. *Psychological Review*, 1966, **73**, 87–103.

LOVEJOY, E. *Attention in discrimination learning.* San Francisco: Holden-Day, 1968.

LUCE, R. D. *Individual choice behavior: A theoretical analysis.* New York: Wiley, 1959.

MACKINTOSH, N. J. Incidental cue learning in rats. *Quarterly Journal of Experimental Psychology*, 1965, **17**, 292–300. (a)

MACKINTOSH, N. J. Selective attention in animal discrimination learning. *Psychological Bulletin*, 1965, **64**, 124–150. (b)

MACKINTOSH, N. J. Comparative psychology of serial reversal and probability learning: Rats, birds, and fish. In R. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning.* London: Academic Press, 1969.

MILES, C. G., & JENKINS, H. M. Overshadowing and blocking in discrimination operant conditioning. Paper presented at the meeting of the Psychonomic Society, Chicago, November 1965.

SHEPP, B. E., & EIMAS, P. D. Intradimensional and extradimensional shifts in the rat. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 375–361.

SHEPP, B. E., & GRAY, V. A. Some effects of variable-within and variable-between irrelevant stimuli on dimensional learning and transfer. *Journal of Experimental Psychology*, 1971, **89**, 32–39.

SHEPP, B. E., & TURRISI, F. D. Learning and tranfers of mediating responses in discriminative learning. In N. R. Ellis (Ed.), *International review of research in mental retardation.* Vol. 2. New York: Academic Press, 1966.

SHEPP, B. E., & TURRISI, F. D. Effects of training procedures on transfer of observing and instrumental responses in discriminative reversals. *Journal of Comparative and Physiological Psychology*, 1969, **69**, 522–527. (a)

SHEPP, B. E., & TURRISI, F. D. Effects of overtraining on the acquisition of intradimensional and extradimensional shifts. *Journal of Experimental Psychology*, 1969, **82**, 46–51. (b)

SHEPP, B. E., & ZEAMAN, D. Discrimination learning of size and brightness by retardates. *Journal of Comparative and Physiological Psychology*, 1966, **62**, 55–59.

STOLLNITZ, F. Can monkeys attend selectively to dimensions of color? Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1969.

SUTHERLAND, N. S. Partial reinforcement and the breadth of learning. *Quarterly Journal of Experimental Psychology*, 1966, **18**, 289–302.

SUTHERLAND, N. S., & HOLGATE, V. Two-cue discrimination learning in rats. *Journal of Comparative and Physiological Psychology*, 1966, **61**, 198–207.

SUTHERLAND, N. S., & MACKINTOSH, N. J. *Mechanisms of animal discrimination learning.* New York: Academic Press, 1971.

TERRACE, H. S. Errorless transfer of a discrimination across two continua. *Journal of the Experimental Analysis of Behavior*, 1963, **6**, 223–232.

THOMAS, D. R. The use of operant conditioning techniques to investigate perceptual processes in animals. In R. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning.* London: Academic Press, 1969.

TRABASSO, T., & BOWER, G. H. *Attention in learning.* New York: Wiley, 1968.

WARREN, J. M., & MCGONIGLE, B. Attention theory and discrimination learning. In R. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning.* London: Academic Press, 1969.

WOLFF, J. L. Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin*, 1967, **68**, 369–408.

ZEAMAN, D., & HOUSE, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency.* New York: McGraw-Hill, 1963.

(Received September 14, 1971)

Psychological Review
1972, Vol. 79, No. 4, 329–343

A CHRONOMETRIC ANALYSIS OF SIMPLE ADDITION[1]

GUY J. GROEN[2] AND JOHN M. PARKMAN[3]

Carnegie-Mellon University

A number of models are considered that specify how children and adults solve single-digit addition problems. It is shown that the most adequate of these in accounting for children's response latencies is a model that assumes the existence of a counter with two operations: setting and incrementing. The child adds two digits, m and n, by setting this counter to max (m,n) and then incrementing it min(m,n) times. This model also accounts for adults' latencies, though with a drastically reduced incrementing time. Some theoretical issues raised by this reduced time are considered, and an alternative model is suggested which assumes that adults usually use a memory look-up process with homogeneous retrieval times, but occasionally revert back to the counting process used by children.

In recent years, it has become common to make a distinction between two kinds of processes in long-term memory: reproductive processes, which are concerned with the retrieval of stored facts, and reconstructive processes, which are concerned with the generation of facts on the basis of stored rules. A principal reason for the emergence of these reconstructive theories is the fact that reproductive theories do not appear to provide a satisfactory explanation of the kind of everyday remembering studied, for example, by Bartlett (1932). Another is that constructive processes appear to play an important role in perception and language. Yet another is the implicit role of constructive processes in Piagetian theory. Unfortunately, most accounts of reconstructive processes have been somewhat metaphorical in their nature and have lacked the explicitness of current models of semantic memory based primarily on reproductive notions. A possible reason for this lack of explicitness, which is shared by many reproductive theories, has been suggested by Reitman (1970) who points out that most memory theories have been concerned with defining global functions rather than specific processes. Since memory may actually consist of a large number of heterogeneous processes, it may be more profitable to investigate the processes applicable to a highly limited task domain. The purpose of this paper is to examine the processes being used in one such limited task domain: the addition of two integers.

The addition of two single-digit numbers can be viewed as either a reproductive or a reconstructive process. Most frequently, it is viewed as the former, consisting of the retrieval of a stored number fact which has been previously learned as part of the addition table. Indeed, this assumption, which goes back at least as far as Thorndike (1922), forms one of the chief bases for the way arithmetic is taught, a child's ability to add two digits being viewed as evidence that he has memorized an association between the two digits and their sum. Although this assumption has never been examined empirically, there are commonsense grounds for rendering such a reproductive process plausible. For example, considerable time is devoted in elementary school to making children memorize the addition table, by means of such techniques as flash cards. Moreover, no effect is usually made to teach any specific algorithmic rule, such as is later taught in connection with the addition of two-digit numbers.

---

[1] Preparation of this paper was supported by Public Health Service Research Grant MH-07722 from the National Institute of Mental Health. The authors would like to thank W. Chase for many helpful comments.

[2] Requests for reprints should be sent to Guy J. Groen, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

[3] Now at the State University of New York at Buffalo.

On the other hand, it is clear that addition is closely associated with some kind of counting process. Although a counting algorithm is almost never specifically taught,[4] addition is usually introduced by demonstrating, in some way, its relationship to counting. Also, children (and some adults) frequently add by overtly "counting on their fingers," despite the fact that behavior of this kind tends to be severely discouraged in the elementary school. In fact, many children actually learn to add by counting before they ever reach the elementary school. Clearly, adding by counting involves the application of a learned rule rather than the retrieval of a number fact. It is a reconstructive process.

There are two advantages to beginning an analysis of the processes used by individuals in solving simple addition problems with a consideration of possible counting models. The first is that given the current state of our knowledge of semantic memory, the possible range of counting models is considerably narrower than the possible range of memory models. The second is that it is a relatively simple matter to construct counting models that will predict systematic differences in reaction times to different problems. This is important because reaction time is the simplest dependent measure to use with problems of this type. Solutions to such problems are usually generated so rapidly that problem-solving protocols yield little insight into the process. Similarly, errors are infrequent enough, except with very young children, that adequate error analysis is not feasible.

A classical technique for inferring, from reaction time data, the nature of the processes involved in performing a simple task is the method of chronometric analysis, first proposed by Donders and more recently reformulated by Sternberg (1969a). Although this technique lapsed into disuse for many years, it has recently received a number of applications to the analysis of processes in pattern recognition (Posner, 1969), memory search (Sternberg, 1969b), and elementary deductive reasoning (Chase & Clark, 1971). It essentially consists of a set of techniques for partitioning reaction times into additive components. The simplest case occurs when a process can be decomposed into a number of identical steps. If this is the case, then it can easily be shown that under suitable statistical assumptions, the reaction time will be a linear function of the number of steps required to perform the task. The memory search processes considered by Sternberg are of this type, as are the counting models considered in the present paper.

FIVE COUNTING MODELS

For problems of the form $m+n$, where m and n are nonnegative integers, and the sum is less than or equal to 9, it is possible to distinguish between five different counting procedures that might be used to solve the problem. In order to distinguish between these procedures, it is convenient to consider a counter on which two operations are possible. One operation is setting the counter to a specified value. It is assumed that this automatically erases the previous value of the counter. The other operation is incrementing the value of the counter by one. The addition operation is performed by setting the counter to an appropriate number and then successively incrementing it by one a suitable number of times. This process is illustrated in Figure 1.

Using this counter, an addition problem can be solved in the following ways:

1. The counter is set to zero; both m and n are added in by increments of one.
2. The counter is set to m (the leftmost number); n (the rightmost number) is then added by increments of one.

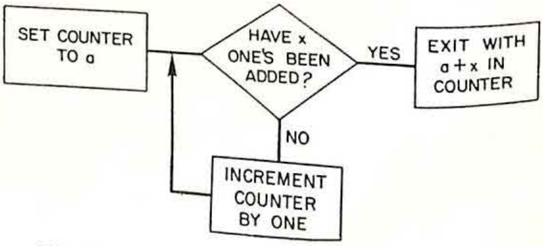


FIG. 1. A device which sets a counter to a and adds x to a by increments of one.

---

[4] An exception is the educational television program "Sesame Street."

3. The counter is set to n (the rightmost number); m (the leftmost number) is then added by increments of one.
4. The counter is set to the minimum of m and n. The maximum of m and n is then added by increments of one.
5. The counter is set to the maximum of m and n. The minimum of m and n is then added by increments of one.

In order to derive predictions from these five models, it is necessary to make two assumptions regarding the time required to perform the setting and incrementing operations: (*a*) The time required to set the counter is assumed to be a random variable A that is independent of the value to which the counter is set. (*b*) The time required to increment the counter by one is assumed to be a random variable B that is independent of the number of times the counter is incremented. Suppose that the counter is set to a certain value and then incremented x times by one. Then the total time T taken by the counter to perform these operations is given by the equation:

$$T = A + Bx \qquad [1]$$

If the counter is being used to solve an addition problem then, corresponding to the five models defined above, x is determined as follows:

Model 1. $x = m + n$
Model 2. $x = n$
Model 3. $x = m$
Model 4. $x = \max(m, n)$
Model 5. $x = \min(m, n)$

For practical purposes, it is inconvenient to consider Equation 1 as it stands, since in order to treat the random variable T directly, it would be necessary to explicitly define the distributions of A and B. An alternative approach is to consider the mean reaction time. Let a be the mean setting time, b the mean time required to increment the counter by one, and let t be the mean reaction time for a given problem. Since x is independent of B, it follows from Equation 1 and the properties of statistical expectations that

$$t = a + bx \qquad [2]$$

where x is determined as before. This is a linear model with two free parameters, a and b. It predicts that given the values of a and b (which are constant over problems), the mean reaction time is completely determined by the value of x, the number of times the counter must be incremented. For a given model, this value can, in its turn, be completely determined for each addition problem by the values of the two numbers being added. As a result, each model can be evaluated by fitting a regression line to observed response latencies. It should be noted that Equation *2* is only intended to apply to correct responses. There are no grounds for assuming that a subject who makes an error is using any of the five counting processes outlined here. It should also be noted that it has only been assumed that the stages of the addition process are additive. Since it has not been assumed that they are statistically independent, it is not necessarily true that var (T) = var (A) + x^2 var (B) (Sternberg, 1969a).

The description of the various counting algorithms has left a number of features undefined. For example, a process that executed these algorithms would need to contain at least one additional feature: some means of keeping track of the number of ones that have been added. The most obvious way to do this is to augment Figure 1 by introducing a second register which is initially set to zero, and is incremented by one each time the main counter is incremented. The test, "Have x ones been added?" would then be performed on this auxiliary register. Since the operation of this register can be assumed to influence each of the five counting algorithms in a uniform fashion, and since it is difficult to conceive of a way in which this operation could affect the linearity of the predictions, it is not discussed any further.

Another feature left undefined, up to now, is the process used, in Models 4 and 5, to find the minimum and maximum addends. Some valuable empirical information is available in this connection. Moyer and Landauer (1967) performed an experiment in which subjects were required to judge which of two numbers was the larger, every possible pairing of the digits 1 through 9 being used. They found that the latency of this judgment increased as the absolute difference beween these numbers decreased. On the surface, such a result might be expected to distort the linearity of Equation 2 if Models 4 or 5 were actually being used. However, given the experimental design used by Moyer and Landauer, the absolute difference and the magnitude of the smallest of a pair of numbers are highly correlated. Parkman (1971) replicated Moyer and Landauer's procedure and found that there was indeed a strong linear relationship between the judgment latency and the magnitude of the smallest of the two numbers in the stimulus display. In fact, this variable turned out to provide a much better fit of the data than the absolute difference originally proposed by Moyer and Landauer. This result suggests a process, quite compatible with the counting algorithms proposed in the present paper, in which the subject searches (or generates) the discrete number line, beginning with one and continuing in increments of one until he finds one of the two numbers. This number is then tagged as the minimum, while the other number is tagged as the maximum. As far as the effect of this process on the reaction times to addition problems is concerned, it is important to observe that the regression line obtained by Parkman, using adult subjects, had a slope of only 10 milliseconds (slopes of about this same order of magnitude have also been obtained in a similar experiment with 10-year-old children). As a result, it seems reasonable to assume that if the regression estimate of the parameter b in Equation 2 is sufficiently greater than 10 milliseconds, then the process of finding the largest and smallest addends has only a second-order effect on the reaction times to addition problems.

Finally, it is clear that there are other undefined processes involved both in perceiving the problem and in making the response. A convenient working assumption is that the effect of these processes on the response latency is a constant. Although this assumption has no effect on either the linearity of the predictions or the value of the slope, it does inflate the intercept. Hence, the empirical estimate of a obtained from the regression analysis cannot be viewed as an accurate estimate of the setting time.

ADDITION IN CHILDREN

While a number of studies of classroom addition have been performed (e.g., Knight & Behrens, 1928), these have all been normative studies designed to determine which problems tended to be most difficult for children. The interpretation of these studies for the purpose of inferring underlying processes is almost impossible because the data were reported in the form of indexes that were composites of various response measures. The only evidence for or against any of the models considered in the preceding section comes in a paper by Wheeler (1939), who showed that the rank ordering of addition problems, on the basis of Knight and Behrens' difficulty index (which was partially based on latencies), was correlated with the smallest of the two numbers being added.

Suppes and Groen (1967) analyzed data from an experiment originally designed for quite a different purpose (comparing addition with subtraction), and found that Model 5 gave a considerably better fit than any of the other four counting models. The subjects in this experiment were children in the first grade of elementary school. Each subject was shown a sequence of addition problems whose sum was less than five, and was required to answer each problem by pressing an appropriate button on a response panel. A regression analysis established that Model 5 gave the best fit to the correct response times. The results of this experiment, as far as Model 5 is concerned, are summarized in Table 1.

It is clear that the means predicted by the regression line are close to the observed means. However, there exist two serious difficulties in interpreting these data. One is that the ties (problems with identical values of m and n) are systematically lower than the predicted values. Indeed, there is no evidence that they exhibit any linear trend at all. The second difficulty is that only two of the problems used in the experiment had a minimum addend of 2. Hence the fact that Model 5 fits the data may merely imply that it is faster to add 0 than to add 1.

An Experimental Test of the Five Models[5]

This experiment was designed to evaluate the five counting models in a more adequate fashion. In particular, a wider range of problems was used. It also seemed desirable to provide a means of deciding whether or not subjects were using different processes. As a result, the experiment was designed so that each subject's data could be analyzed separately.

Method. The subjects consisted of 37 children (15 males and 22 females) in the first grade of an elementary school in a predominantly upper-middle-class community in the San Francisco Bay area. Their average age was 6 years 10 months, and their average Binet IQ was 125. The experiment was conducted toward the end of the school year, and all the subjects had been taught addition. Enquiries of the teachers revealed that none of the counting algorithms had been specifically taught.

The materials used as stimuli consisted of all 55 problems with a sum less than or equal to 9. The problems were presented as row-addition problems of the form $m + n = —$, the subject's task being to find the missing number. Six different random presentation sequences, each containing exactly the same set of 55 problems, were prepared. The random ordering was constrained in such a way that the following sequences did not appear: two problems with the same answer; $m + n$ followed by $m + k$; $m + n$ followed by $k + n$.

The subject saw each problem projected onto a screen by means of a Carousel Projector. He had in front of him a response panel with a row of 10 buttons marked 0 through 9, and made his response by pressing one of these buttons. He was allowed as long as necessary to make the response, and there was an interval of $2\frac{1}{2}$ seconds between the subject's response and the presentation of the next problem. During this interval, the correct answer was displayed on the screen. Both the slide projector and the response panel were connected to a Beckman timer, which measured the response latency to an accuracy of 5 milliseconds.

The experiment was performed in a soundproof trailer located on the premises of the school. Subjects went from their classes to the experiment in the course of a normal school day. Each subject was run in sessions lasting at most 15 minutes each on five successive days.

TABLE 1

REACTION TIMES OF CHILDREN TO ADDITION PROBLEMS

Min(m,n)	Number of observations	Reaction times		
		Predicted	Observed	Ties
0	10	3.26	3.23	2.98
1	7	3.97	4.05	2.67
2	2	4.68	4.52	3.70

The first day began with a preliminary task, in which the subject was successively shown the numbers from 0 to 9 in random order. His instructions were to press the button that was the same as the number he saw on the screen. After having completed this preliminary task, the subject was told that he would be shown some slides on which there were two numbers, and that he was to "press down on the button that is the same as the answer you get when you add the two numbers together." The main sequence of 55 addition problems was then presented. During the course of the experiment, the experimenter sat behind the subject, so that he could not be seen unless the subject deliberately turned around. If the subject seemed at all uncertain of what to do, he was told to punch the button that was the same as the missing number.

The procedure during subsequent sessions was almost identical to that on the first day, the subject being required to perform both the preliminary task and the main task. The only differences were that the instructions were omitted and a different sequence of problems was used on each day, according to a suitably counterbalanced design.

Results. The first day of the experiment was viewed as a training day, the function of which was to familiarize the subjects with the apparatus and with the task. The evaluation of the models was based on the mean latency of a correct response, computed over the last four days for each subject and each problem. Hence, each data point was a mean based ideally on four observations. In computing these means, two types of observations were omitted: incorrect responses and outlines with excessively long response latencies. As far as the outliers were concerned, the cutoff point was set at 10 seconds. All responses with a latency greater than this value were omitted from the analysis. The average rate of outliers was under 1%. Only 12 subjects had any outliers at all, and only 2 had more than three outliers. The overall error rate was also low. No subject had an error rate of more than 10%, the mean being 3%. Twenty-seven subjects had an error rate of less than 5%. The problems with the highest error rates were $2 + 5$, $5 + 2$, $2 + 6$, and $5 + 3$. These four problems all had 13 errors out of a possible 148, an error rate of slightly under 9%.

---

[5] A more extended account of this experiment is given by Groen (1967).

The models were evaluated by fitting five regression lines, one for each model, to the mean success latencies of each subject over the last four days. In this way, five values of F, the standard statistic for the significance of the slope parameter b, were obtained for each subject. In order to decide which of the five models provided the most adequate account of the data, two criteria were used: The first criterion was which model maximized F. This criterion was used to select the best model for a given subject. The second criterion was whether $F(df = 1/53)$ was significant at the .01 level. This criterion was used to reject subjects whose maximal F might reflect random error rather than the fit of any model.

The outcome of applying this analysis to the data of each subject was that an individual's data were either accounted for most adequately by Model 5 or they were not adequately accounted for by any of five models. Twenty out of the 37 subjects received significant maximum Fs on Model 5. All but one of the remaining subjects failed to receive a significant F on any of the five models.

It seemed possible that the superiority of Model 5 over Model 1 might have been due to a statistical artifact, since the independent variable had a range of 5 values for this model, whereas it had a range of 10 values for Model 1. In order to check this possibility the entire analysis was repeated using a subset of 18 problems selected in such a way that Model 1 placed six of them at each of the points $x = 7$, $x = 8$, $x = 9$, while Model 5 placed six at each of the points $x = 1$, $x = 2$, $x = 3$. The results of this analysis were almost identical to the preceding one, except for a tendency for the values of F to be lower.

For subjects whose data were best accounted for by Model 5, the value of the intercept parameter a ranged from 1.87 seconds to 4.65 seconds, while the slope parameter b ranged from .12 second to .73 second. In order to obtain an overall estimate of these parameters, the mean success latencies, averaged over the 19 subjects best fitted by Model 5, were computed for each problem. A regression line was then fitted for each model. Not surprisingly, Model 5 gave by far the best fit ($F = 70.8$ vs. $F = 10.6$ for Model 1). The intercept parameter was estimated to be 2.53 seconds while the slope was .34 second. These estimates are sufficiently close to those obtained by Suppes and Groen (1967) to indicate the results of the two experiments are actually reflecting the same process. The estimate of the slope is also close to the estimate of the speed of silent counting (Beckwith & Restle, 1966; Landauer, 1962), which adds a considerable amount of plausibility to the notion that subjects are, indeed, using an actual counting process as defined by Model 5.

Other Processes

While this analysis shows that Model 5 gives a better account of the data than any of the other counting models, it does not indicate how well the model fits in an absolute sense. That this fit is far from perfect is implied by the fact that the value of R^2 (the proportion of the variance accounted for by the regression line), computed for the means averaged over subjects, is only .57.

To examine the actual goodness of fit, the mean success latency to each problem was plotted as a function of the minimum addend. The resulting plot, with the mean success latency rounded to the nearest tenth of a second, is shown in Figure 2. In this graph, a solid dot indicates an observed mean response latency. The figures beside it denote the problems whose mean success latency had the plotted value. Thus, mn is used to denote $m + n$. For example,

• 70, 90

signifies that the problems $7 + 0$ and $9 + 0$ both had a mean success latency specified by the •. In the figure, $7 + 0$ and $9 + 0$

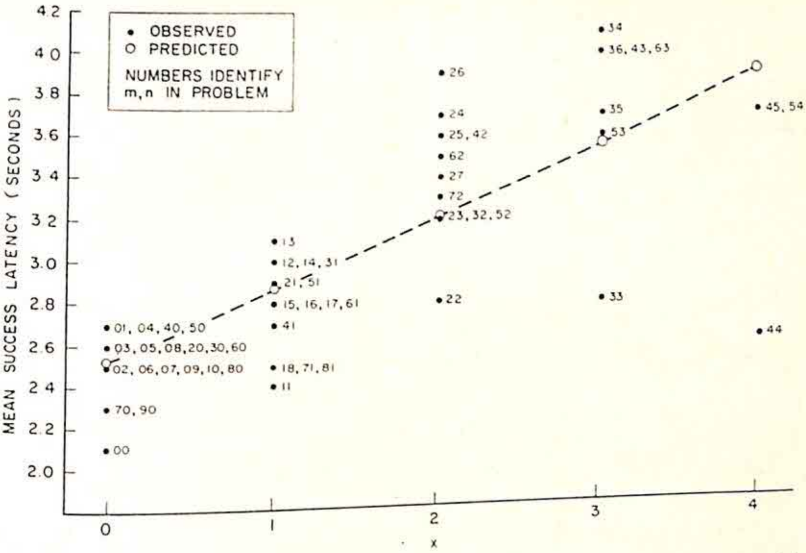


FIG. 2. Overall mean latencies to each problem as a function of the minimum addend. (Dashed line indicates regression line of Model 5.)

both had a latency of 2.3 seconds. The white dots denote the values predicted from the regression line for each x. The dashed line is the regression line resulting from the analysis just described.

It is clear from Figure 2 that the most important sources of lack of fit are the ties. Essentially the same phenomenon occurs in this experiment as occurred in the data of Suppes and Groen (1967). The ties always have a lower latency than any of the other problems with the same minimum addend. In the case of $2 + 2$, $3 + 3$, and $4 + 4$, this discrepancy is quite considerable.

In order to account for these excessively short mean success latencies it is convenient to assume that they are faster then they should be because the subject has memorized the correct response. In other words, it could be said that an association had been formed between the stimulus of the problem and the response consisting of the correct solution. Of course, this need not be a "pure associative connection" in the Thorndikean sense. It might be that the subject has the response stored in long-term memory and uses some type of fast retrieval algorithm in order to find the correct response. The important point is that it is assumed that whatever process is responsible for retrieving the correct response takes a length of time that is short compared to the length of time required to generate the same answer by means of a computational procedure, and is independent of the magnitude of the number being retrieved.

To be explicit, suppose that the solutions to ties have been memorized. Then, for problem of the form $m + n$, Equation 1 can be changed to:

$$T = C \text{ for } m = n, \qquad T = A + Bx \text{ otherwise,} \qquad [3]$$

where C is a random variable representing the length of time required to retrieve the solution from memory, and A and B have the same meaning as in Equation 1. Note that C is assumed to be independent of the problem, since it seems reasonable to assume that differences in retrieval time as a function of the problem are small compared to random variations in C. Much the same type of consideration lay behind the notion of making the setting time A independent of the problem. In fact, A and C are similar in that they cannot be viewed as "pure" setting times and retrieval times, since both are composites of sensory, perceptual, and motor factors as well as whatever internal process corresponds to setting and retrieving.

Equation 3 could be directly transformed into expectations and fitted to the data as was done with Equation 1. However, this procedure has the disadvantage of introducing an extra parameter $E(C)$ into the models. As a result, it would be extremely difficult to decide whether an increased goodness of fit resulted from the introduction of a memory assumption or merely from the introduction of an extra parameter. To avoid this difficulty, it will be assumed that

$$E(C) = E(A) = a.$$

Equation 3 then becomes

$$t = a \text{ if } m = n,$$
$$t = a + bx \text{ otherwise.} \qquad [4]$$

In order to decide whether or not individuals actually tended to use a process of the sort described for the ties, the regression analysis, in which each model was defined by Equation 2, was repeated, with each model being defined by Equation 4. Essentially, all this involved was setting the value of x to 0 for the ties. The same degrees of freedom were used in both analyses. A comparison between the old and new analyses showed an increase in the value of F obtained on Model 5 for 18 out of the 20 subjects originally best fitted by the old version. Moreover, 5 of the 16 subjects who had failed to obtain a significant F on any of the old models obtained a significant maximum F on the new Model 5. These increases were not accompanied by increases in any of the other four models. The analysis was also repeated on the overall problem means of Figure 2. The new Model 5 was by far the best fitting model ($F = 166$, $R^2 = .76$), with an increased slope (.41 second) but an unchanged intercept (2.51 seconds). For individual subjects, the values of R^2 ranged from .18 to .76 with the new model, as compared to a range of between .12 and .52 for the old model. While this new Model 5 produces a much better fit, it still leaves about 25% of the variance unaccounted for.

One possible source of this error variance may be the fact that a keyboard was used to record the responses, since the mean reaction times for each key (computed from the preliminary task) did show a pronounced serial position effect. On the whole, the keyboard responses could be divided into groups with approximately equal times: 0–2 (1.65 seconds), 3–5 (1.85 seconds), and 7–9 (1.80 seconds). While the difference between the first group and the others is appreciable, only 4 of the 55 problems involved responses in this group. The difference between the remaining two groups is small compared to the magnitude of the slope parameter.

A more interesting reason for the remaining lack of fit can best be seen from an inspection of Figure 3, which shows the regression line resulting from the modified Model 5, together with the mean observed latency for each minimum addend (ties are assumed to have a minimum of zero). A comparison of Figure 3 with Figure 2 reveals that although the fit of the regression line to most problems is greatly improved, the fit to problems with a minimum addend of four is somewhat worse than before. The most likely explanation is that some subjects have the solutions to these problems stored in the same kind of fast access memory as has been proposed for the ties. In fact, a distributional analysis shows the frequency distribution of latencies to these problems to be bimodal. Moreover, it turns out to be possible to assign each subject in an unambiguous fashion to one of two distribu-

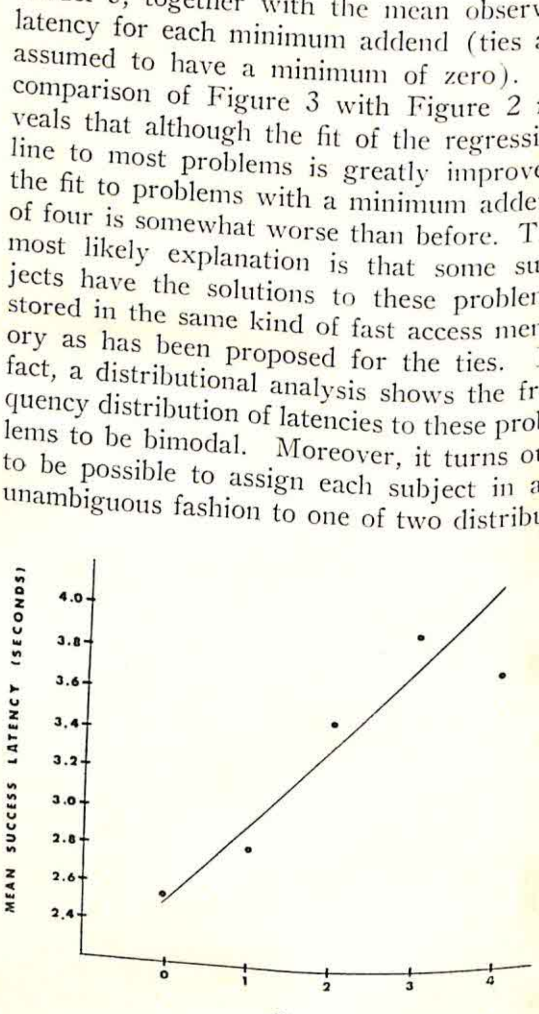


FIG. 3. Mean latencies, averaged over all problems with the same minimum addend (ties assigned minimum addend of zero). (Solid line indicates regression line of modified Model 5.)

tions (Groen, 1967). The mean of the slower group is 4.5 seconds, which is very close to the value predicted by the regression line of Figure 3, whereas the mean of the faster group is 2.3 seconds, which is close to the mean latency for the ties.

This result raises the question of why ties and problems with a minimum addend of four shoud be stored in a fast access memory whereas other problems are not. The most plausible hypothesis is that it is only possible to store the solution to a problem in fast access memory, and be able to retrieve it, if the subject can find a distinctive feature which can be used as an "address" for the solution. In other words, the subject must be able to discriminate a property of the problem which serves to differentiate it from problems with different solutions. In the case of ties, such properties as "all ones" are readily apparent in the visual display and should be equally salient for all subjects. The distinctive features of $5 + 4$ and $4 + 5$ arise as an artifact of the experimental design: they are the only problems with a minimum addend of four, and they both have the same solution. This is clearly a more subtle cue, which subjects can be presumed to learn to use as they become familiar with the experiment. Since each subject attended only five experimental sessions, it is to be expected that some subjects would not yet have learned to utilize the cue. However, the notion that subjects utilize it at all is consistent with the notion that subjects are utilizing an algorithm based on the minimum addend in solving other problems.

ADDITION IN ADULTS

By the time a person has reached adulthood, addition has become a highly overlearned skill. The purpose of this section is to examine what changes occur as a result of this overlearning. The preceding analysis suggests two possibilities. The first is that there is no change in the process used, except possibly for a decrease in the incrementing time and the setting time. The second possibility is that individuals gradually cease to use the process defined by Model 5. Instead, they memorize the answer, and it is stored in a fast access memory. This mechanism was invoked in the discussion of children's addition in order to account for the unusually low response latencies of ties. It can be assumed that ties are the first problems to be memorized in such a way because they are easily encodable. As a result of overlearning, the individual might develop efficient ways of encoding the other problems.

There are other possibilities too. In particular, there are the other counting models suggested in the context of children's addition. Also, two theories have been proposed specifically in the context of adult addition. Thomas (1963) has suggested that the mental processes underlying simple calculation behave as an information-processing system for which the time needed to perform a given arithmetic operation is strictly proportional to the amount of information that must be handled by the system. From this assumption, and some notions about how the components of a problem may be combined in terms of total information, he has derived the prediction that reaction times to simple addition problems should be a linear function of $\log\ (m + n)$.

Restle (1970) has tentatively proposed a model which views the processes underlying addition as analogue operations, in contrast to the digital operations assumed by the counting models. According to this model, numbers are represented internally in the form of number lines. The individual performs an addition by converting the magnitudes of m and n into disjoint segments of a continuous internal number line and then estimating the length of the two joined line segments. Restle suggests that when m and n differ to increasing extents, mental manipulations of the distances postulated by this model should be easier, and latencies should decrease. While the form of the function is unspecified, the prediction that response latencies should be a linear function of $|m - n|$ (which Restle terms the split between the two numbers) can be viewed as a first approximation.

In principle, it is simple to test between these alternative possibilities. The first implies a straightforward linear fit of Model 5. Both Thomas's model and the linearized

TABLE 2

COMPARISON OF R^2 FOR VARIOUS MODELS

Model	Independent variable	Value of R^2	
		Adults	Children
1	$m + n$	.71	.11
2	m	.18	.01
3	n	.36	.04
4	$\max(m,n)$	.34	.02
5	$\min(m,n)$	.73	.80
Thomas	$\log(m+n)$	.58	.12
Restle	$\lvert m-n \rvert$	.09	.32

Note.—Adult data from Parkman and Groen (1971).

Restle model imply equally straightforward linear fits of log $(m + n)$ and $|m - n|$, respectively. On the other hand, the fast access memory model implies the existence of no systematic effects at all. In particular, it implies no systematic differences between ties and other problems. However, because the response latencies of adults to problems of this kind are much faster than those of children, it is difficult to investigate these processes by means of the simple experimental paradigm used with the children. The production of a straightforward solution to an addition problem requires a voice key or a keyboard to record the latencies. Apparatus of this kind tends to produce extraneous effects that may mask small, but systematic, latency differences. As a result, it is necessary to make use of a more indirect paradigm with simpler responses.

Such an indirect paradigm was adopted by Parkman and Groen (1971), who showed college students a series of problems of the form $m + n = k$, where m and n were single digits while k was a pair of digits. The subjects task was to press a yes button if k was equal to $m + n$ (e.g., $5 + 4 = 09$) and to press a no button if k was unequal to $m + n$ (e.g., $5 + 4 = 11$). On half of the trials (positive instances), the presented solution was equal to the true sum. On the remaining trials (negative instances), the presented sum differed from the true sum by not more than ±2. Apart from this, the procedure was the same as with the children, except that a broader range of problems was used (all problems with a true sum of less than 18).

On the basis of the data from this experiment, a value of R^2 can be computed for each model by fitting the appropriate regression line to the mean success latencies of individual problems. These values are shown in Table 2. In this analysis, ties have been omitted, and positive and negative instances have been pooled. For purposes of comparison, the corresponding values of R^2 obtained from the children's data, again omitting ties, are also shown. It can be seen that the adult data resemble the children's data in that the minimum addend (hereafter referred to as the min) provides the best fit. It differs in that the sum of the two addends (hereafter referred to as the sum) accounts for almost as much variance as the min, whereas this is clearly not the case in the children's data. The fact that the statistic suggested by Thomas accounts for 58% of the variance of the adult data can be attributed to the obvious correlation of this variable with the sum. The moderately high values of R^2 obtained by the max, in the case of the adult data, and Restle's split, in the case of the children's data, reflects residual effects that are small compared to the effects of either the min or the sum.

Since the task used with adults was not a pure addition task, the results of Table 2 are not sufficient, by themselves, to reject any of the proposed models. However, it will be most convenient to put this question aside for the moment and concentrate on the effect of a single variable: the min. Figure 4 shows the mean success latencies of positive and negative instances as a function of this variable, together with the appropriate regression lines. Ties are also shown, plotted separately from nontie problems, with positive and negative instances pooled. This graph has two main features. The first is that ties have the same characteristics as with children: they are uniformly lower than nontie problems and show no linear trend as a function of either the min or any of the other variables of Table 2. The second main feature is that the regression lines for positive and negative in-

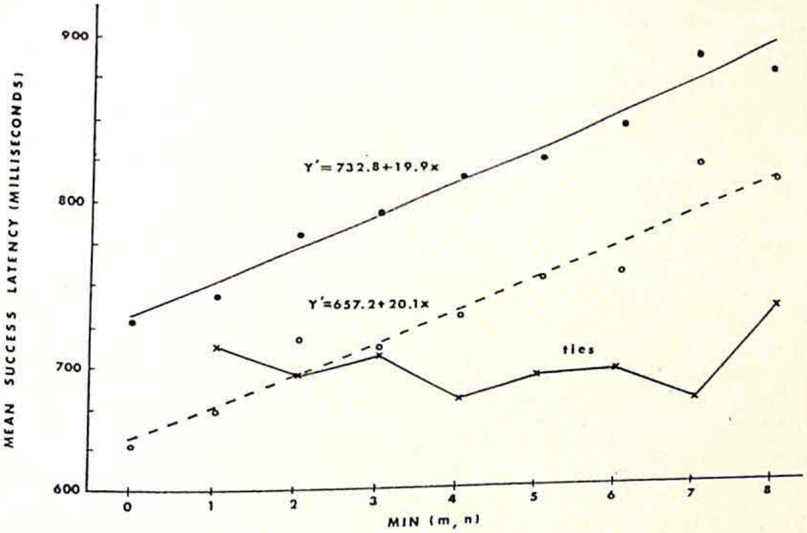


Fig. 4. Mean latencies of adults as a function of the minimum addend. (Filled circles refer to negative instances, crosses refer to ties. Data recomputed from Parkman & Groen, 1971.)

stances have different intercepts (657 milliseconds for positive instances and 733 milliseconds for negative instances) but identical slopes (20 milliseconds). This finding that the lines for positive and negative instances are parallel is evidence that the experimental variables determining the slope are affecting a different stage than are the variables determining the intercepts, and suggests that the processing durations of these respective stages are related in an additive manner (Sternberg, 1969a). This implies that the process used to solve these problems can be viewed as having two additive stages: one which has a duration proportional to the min, and another with a duration that depends on the type of instances that have been presented.

In order to discuss the nature of this process more precisely, it is convenient to introduce the notion of a salient structural variable. A structural variable is a variable, such as the min, whose value can be determined from the structural components of the problem. Such a variable is salient if it accounts for the data, either by itself or in combination with other variables. Since a structural variable is defined in a process-independent fashion, it corresponds to a class of processes rather than to a unique process. Hence, from an extremely strict point of view, all that can be inferred from the results of Figure 4 is the existence of an additive combination of two salient structural variables: the min and the type of instance. To stop at this point is clearly unsatisfactory. To go further, however, it is necessary to use arguments based on plausibility rather than on strict necessity. In this connection, a distinction can be made between two kinds of plausibility: empirical and theoretical. A process is empirically plausible if it is sufficient to account for the salient variables. It is theoretically plausible to the extent that it can be defined in terms of a well-defined model of minimal complexity.

The most plausible first step in explicating the processes underlying the results of Figure 4 is to propose that the min corresponds to an addition stage, during which the sum of *m* and *n* is somehow found, while the type of instance corresponds to a second stage during which the result of the addition stage is used to generate a yes or a no response. There are a number of arguments in favor of this proposal. One is the fact that the min turned out to be the salient structural variable in the experiment with children, which made use of straightforward addition problems. Another is that it seems currently impossible to postulate a theoretically plausible process that makes the min reflect anything other than addition. Finally, it is possible to give a quite detailed account of the processes underlying the second stage (Parkman & Groen, 1971). Since the main issues at stake in the present paper are concerned with the addition stage, the processes underlying the second stage are not described here, except to mention in passing that the duration can be accounted for by either a comparison process or a response organization process.

A similar analysis can be made in terms of the other structural variables. When regression analyses are performed on positive and negative instances separately for a given variable, both the values of R^2 and the estimates of the slope parameters are identical. This implies that a low value of R^2 in Table 2 actually does imply a relative inadequacy of a given model in accounting for the addition process. This can be interpreted as providing fairly decisive evidence against Thomas's theory as far as this range of problems is concerned. The same is not true of Restle's, since only a linearization of a possibly nonlinear model was tested. It can be concluded, however, that if the process suggested by Restle is actually used by subjects in solving addition problems, then that process cannot possibly be linear for the range of addends considered in this experiment.[6]

It might seem legitimate to conclude that the process used by adults is identical to that used by children. However, there are two serious obstacles to a conclusion of this kind. One is the fact that the sum is as salient a structural variable as the min (regression analyses yielded intercepts of 616 milliseconds for positive and 688 milliseconds for negative instances, and slopes of 11 milliseconds in both cases). There are a number of explanations of this result. One is that adults use the underlying process defined by Model 1. This has the problem of begging the question of why a developmental transition should occur from Model 5. Another possibility is that the underlying process actually is that defined by Model 5, but it differs from the process used by children in certain detailed aspects whose nature cannot be revealed by the approach used in the present paper.

The other obstacle is the fact that the slope of the regression line for children is 400 milliseconds whereas the corresponding slope for adults is 20 milliseconds. Thus, the adult slope differs from the children's slope by a factor of 20. Since the rate of silent counting in adults is at best one number every 150 milliseconds, it is necessary to conclude that if adults actually do use the process defined by Model 5, then they are not incrementing by counting to themselves. In children, on the other hand, the coincidence of the incrementing rate for addition with the speed of silent counting gives some grounds for believing that the process actually is based on silent counting. In other words, if Model 5 holds for both children and adults, then the incrementing process based on counting that is used by children must be replaced by an incrementing process of an unknown nature.

It may not be entirely coincidental that the memory search rate obtained by Sternberg (1969b) is of the same order of magnitude as the presumed incrementing rate for adult addition. The counting process defined by Model 5 can be viewed as a memory search process in which the setting operation corresponds to generating a list of numbers and the incrementing process corresponds to finding the next item on a list. However, further speculation along these lines would be premature. The correspondence between the memory search that might be used in adult addition and the exhaustive search postulated by Sternberg is not exact. Also, one would expect developmental changes in search rates, using Sternberg's paradigm, that corresponded to the change in slope observed in the addition experiments. While the evidence is quite tentative, it currently seems that the rate of memory search for young children is approximately the same as for adults (Hoving, Morin, & Konick, 1970).

---

[6] Particular tasks, especially the mode of response employed, may make a most important difference in the processes subjects use. It is not inconceivable that Restle's task would encourage subjects to use approximate and perhaps analogue methods.

Because of this difficulty in defining the change in process underlying this change in slope, it is desirable to consider a quite different kind of explanation. Essentially, this implies that the observed difference in slopes is due to an artifact of averaging. Suppose that an adult has the solutions to addition problems stored in a fast access memory. Suppose, however, that he occasionally reverts to a childish way of performing addition. More precisely, assume that this childish way of doing addition is defined by Model 5, and that the probability p that such a reversion will take place is uniform over all problems. Using the notation of Equation 3, it follows that

$$T = C \text{ with probability } 1 - p,$$
$$T = A + B \min(m, n) \text{ with probability } p.$$

For simplicity, it can again be assumed that C and A have identical means. The standard definition of conditional expectation then implies that the mean latency t is given by the equation:

$$\begin{aligned} t &= a(1-p) + (a + b \min(m, n))p \\ &= a + pb \min(m, n). \end{aligned}$$

In this equation, the response latency is a linear function of the min. Hence, the predictions made by this new model will be indistinguishable from those made by assuming a change in the rate of incrementing. However, it is now clearly unnecessary to assume an incrementing rate that is faster than the speed of silent counting. Suppose, for example that the rate for adults is actually the same as the rate for children (i.e., about one increment in 400 milliseconds). Then an observed slope of 20 milliseconds would result in a value of p of about .05. In other words, a reversal to the childish method would have to occur only about once in every 20 problems to account for the observed adult slopes.

The assumptions used in deriving this equation have been made as strong as possible in order to avoid obscuring with algebraic complexity the basic point that it is unnecessary to postulate an unknown high-speed counting process in order to account for the linearity of the adult latencies as a function of the min. It should be mentioned, however, that these assumptions can be relaxed in two ways without affecting either the linearity of the min or the prediction of a drastic decreases in slope:

1. It can be assumed that the mean setting time is not equal to the mean retrieval time. This will influence the value of the intercept, but not the slope. It will be recalled however, that the empirical estimate of the intercept reflects additional motor and perceptual processes. It cannot be assumed that these processes are identical in the two experiments, since there are obvious differences both in the age of the subjects and the physical nature of the stimuli and the responses. Hence the observed decrease in the estimate of the intercept that occurs with the adults may reflect changes in this latter class of processes.

2. It can be assumed that the probability p of reverting to the counting process is not uniform over problems. In this case, what ensues depends on the distribution of p. If the distribution is random, then only the standard error of the slope estimate would be affected. On the other hand, systematic variations in p would be reflected in systematic variations in the latencies as a function of some structural variable other than the min. In general, there is no reason why this variable should combine with the min in an additive fashion.

To summarize, the difference in slopes between children and adults can be accounted for in two ways: by a single-process model with an unknown counting process, or by a two-process model in which the reproductive process which children already use in a local fashion to solve ties such as $2 + 2$ is used by adults for all problems most of the time. From a purely formal point of view, there is no way of distinguishing between the two models on the basis of the mean latencies. However, if there is no real correspondence between the postulated high-speed counting process and the high-speed scanning discussed by Sternberg, then the two-process model provides a considerably more plausible explanation of the data than does the single-process model. It might be objected that the two-process model does not explain what causes the strategy switch. Presumably, it would occur if retrieval were to fail. While it is currently impossible to elucidate this process, it is well known from such observations as the "tip of the tongue" phenomenon (Brown & McNeill, 1966) that such failures to retrieve are quite common. It is reasonable to anticipate that a person confronted with such a failure to retrieve would fall back on a reconstructive algorithm.

The main emphasis of this paper has been on reconstructive processes. The reason for this emphasis is that it turns out to be a relatively simple matter to formulate well-defined, intuitively plausible reconstructive models that yield strong empirical predictions. On the other hand, the possibility remains open that the latency phenomena described in this paper may arise from a complex reproductive process. In fact, it is not difficult to think of a number of retrieval models in this vein. For example, the counting models could easily be reformulated as retrieval algorithms that calculated an index, rather than a sum, with the index being used for a memory retrieval operation. Alternatively, one might reformulate each model in terms of a list structure. The setting operation might then correspond to an operation that accessed a given list, while the incrementing operation might correspond to finding the next element on the list. Models of this kind have not been discussed more extensively in the present paper because it has not been found possible, up to now, to formulate a well-defined model along these lines without introducing an excessive number of arbitrary assumptions. Moreover, a variety of reproductive models, based on quite different assumptions, are indistinguishable on the basis of the kind of experiment considered here. Despite this unsatisfactory state of affairs, some indirect support for this possibility comes from an investigation by Parkman (1972) of multiplication in adults. Using a direct analogue to the procedure of Parkman and Groen (1971), it was found that latencies increased monotonically as a function of the minimum multiplicand. However, there are several discrepancies between the results for multiplication and the results for addition. The most important of these arises with the ties. It will be recalled that latencies of ties in addition problems are not related to the size of the addend. In multiplication, the latencies of ties turn out to increase as the size of the multiplicand increases. Any theory that attempted to account for multiplication and addition on the basis of the same reproductive process would have to explain why this relationship is present in multiplication but absent in addition. At present, the task of formulating a satisfactory explanation seems fraught with difficulties. On the whole, the question of precisely which process underlies performance on simple addition problems is likely to remain open until considerably more is known about the role of reconstructive and reproductive processes in other aspects of semantic memory. The present paper merely points to some promising possibilities and some potentially useful experimental results.

REFERENCES

BARTLETT, F. *Remembering.* London: Cambridge University Press, 1932.

BECKWITH, M., & RESTLE, F. Process of enumeration. *Psychological Review,* 1966, **73,** 437-444.

BROWN, R., & MCNEILL, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5,** 325-337.

CHASE, W. G., & CLARK, H. H. Semantics in the perception of verticality. *British Journal of Psychology,* 1971, **62,** 311-326.

GROEN, G. J. *An investigation of some counting algorithms for simple addition problems.* (Tech. Rep. No. 118) Stanford, Calif.: Stanford University Institute for Mathematical Studies in the Social Sciences, 1967.

HOVING, K. L., MORIN, R. E., & KONICK, D. Recognition reaction time and size of the memory set: A developmental study. *Psychonomic Science,* 1970, **21,** 248-249.

KNIGHT, F. B., & BEHRENS, M. *The learning of the 100 addition combinations and the 100 sub-*

traction combinations. New York: Longmans, 1928.

LANDAUER, T. K. Rate of implicit speech. *Perceptual and Motor Skills,* 1962, **15**, 646.

MOYER, R. S., & LANDAUER, T. K. Time required for judgments of numerical inequality. *Nature,* 1967, **215**, 1519–1520.

PARKMAN, J. M. Temporal aspects of digit and letter inequality judgments. *Journal of Experimental Psychology,* 1971, **91**, 191–205.

PARKMAN, J. M. Temporal aspects of simple multiplication and comparison. *Journal of Experimental Psychology,* 1972, in press.

PARKMAN, J. M., & GROEN, G. J. Temporal aspects of simple addition and comparison. *Journal of Experimental Psychology,* 1971, **89**, 333–342.

POSNER, M. Abstraction and the process of recognition. In G. Bower (Ed.), *The psychology of learning and motivation.* Vol. 3. New York: Academic Press, 1969.

REITMAN, W. What does it take to remember? In D. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

RESTLE, F. Speed of adding and comparing numbers. *Journal of Experimental Psychology,* 1970, **83**, 274–278.

STERNBERG, S. The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica,* 1969, **30**, 276–315. (a).

STERNBERG, S. Memory-scanning: Mental processes revealed by reaction time experiments. *American Scientist,* 1969, **57**, 421–457. (b)

SUPPES, P., & GROEN, G. J. Some counting models for first grade performance data on simple addition facts. In J. M. Scandura (Ed.), *Research in mathematics education.* Washington, D. C.: National Council of Teachers of Mathematics, 1967.

THOMAS, H. B. G. Communication theory and the constellation hypothesis of calculation. *Quarterly Journal of Experimental Psychology,* 1963, **15**, 173–191.

THORNDIKE, E. L. *The psychology of arithmetic.* New York: Macmillan, 1922.

WHEELER, L. A comparative study of the difficulty of the 100 addition problems. *Journal of Genetic Psychology,* 1939, **54**, 295–312.

(Received November 9, 1971)

Psychological Review
1972, Vol. 79, No. 4, 344–358

LAWS OF THE SIMPLE VISUAL REACTION TIME[1]

WARREN H. TEICHNER[2] AND MARJORIE J. KREBS

New Mexico State University

Fourteen studies from the literature on the reaction time (RT) to a flash of light were found which provided sufficient methodological detail and data to allow for a quantitative analysis of the effects of the following selected variables: luminance, duration, size of stimulus, response to onset versus termination of the signal, and monocular versus binocular viewing. Mathematical relationships were developed which can be used to predict binocular RTs over a wide range of luminance, signal duration, and signal size. It was also shown that the product of RT and luminance may be used to represent a response criterion in the sense implied by the theory of signal detection as developed in recent latency models. The results were interpreted as suggesting two general energy-processing activities—criterion development and an impulse "countdown"—going on in parallel which are in series with a sensory delay.

The purpose of this study was to develop a quantitative understanding of the manner with which the simple visual reaction time (RT) depends on the physical parameters of the stimulus. A fundamental premise of the method employed was that the absolute values of RT reported in the literature are sufficiently reliable to be used for that purpose. In a sense, this study was also an evaluation of the quantitative generalizability of the RT literature.

Parameters other than those which describe the signal are also important (Teichner, 1954), but have been largely ignored in this study. Since the present data were the absolute RT values of published studies, differences in the data among studies that might be associated with the effects of the ignored variables were assumed to be small and unsystematic. There were three exceptions. First, uncertainty in the time of arrival of the signal was assumed to be important, and only studies for which this uncertainty was thought to be minimal were selected for analysis. Second, positional uncertainty was excluded by using only studies in which the location of the signal was known in advance by the subject. Third, to reduce the scope of the effort, only studies involving foveal vision and signals of white light were considered. Interest was focused on (*a*) the RT function for a flash of white light having no positional uncertainty and little or no temporal uncertainty, (*b*) the possibility of developing quantitative RT models, and (*c*) the question of whether the available studies are sufficiently interconsistent to permit the above two interests to be satisfied. The variables of concern were the luminance, duration, and size of the signal, whether viewing was monocular or binocular, and whether response is faster to the onset or the termination of the light. The data of all studies published between 1886 and 1969, considered acceptable for quantitative use, were converted to common units of measurement and treated as the raw data of this study. Most of the studies which were not used were rejected for failure to provide information about all of the needed signal characteristics.

THEORETICAL BACKGROUND

Theoretical interest in RT has been continuous, at least since Berger (1886) and Cattell (1886) showed that RT is inversely related to the intensity of the visual stimulus.

---

[1] This research was supported by a contract between the Office of Naval Research and New Mexico State University. The authors are indebted to William Mayhall and Dean LeMaster for their assistance, and to Martin Tolcott of the Office of Naval Research for his encouragement. The authors are also grateful to David Raab for his constructive review of this paper and for his suggestions.

[2] Requests for reprints should be sent to Warren H. Teichner, Department of Psychology, New Mexico State University, Box 3452, Las Cruces, New Mexico 88001.

The intensity relationship has since been demonstrated for RT to signals in other sensory modalities and may be considered to be well established, at least qualitatively. Less well described are the effects of other signal characteristics.

Since Piéron (1920), theoretical approaches to the explanation of RT phenomena, especially the intensity relationship, have been concerned with one or more of three postulated kinds of RT components: (*a*) a very short delay required for neural transmission, (*b*) a delay which depends on the intensity or energy characteristics of the signal, and (*c*) as emphasized recently by Grice (1968), a delay which depends on attentional, perceptual, cognitive, and motivational factors. Piéron (1920) proposed a hyperbolic RT function expressed in terms of the first two postulated components noted above:

$$\mathrm{RT} = \frac{a}{I + b} + K, \qquad [1]$$

where a and b are empirical constants; I is stimulus intensity expressed in multiples of threshold intensity, for example, in multiples of threshold luminance for visual signals; and K is an "irreducible minimum" neural delay which is constant, at least for a given sensory modality and a given effector system.

The fraction, $a/(I + b)$, was proposed as the "reducible margin," a quantity which decreases with intensity and is asymptotic to K as a limit. Although Piéron found that Equation 1 could be fitted nicely to a variety of data obtained in his laboratory, as Welford (1960) has noted, the constant, b, in those curve fits was unique to each set of data fitted. While Welford suggested that this may have been due to the difficulty of obtaining high precision data, it is also plausible that either some other function may be more appropriate for providing a general law, or that the general form of the relationship used by Piéron may be appropriate, but may not include all of the critical factors. Since Piéron, a number of investigators have proposed a variety of other functions to express the dependence of RT on stimulus intensity (e.g., Bartlett & McLeod, 1954; Hull, 1949; Vaughan, Costa, & Gilden, 1966; Woodworth & Schlosberg, 1954).

Stochastic models of RT (e.g., Luce & Green, 1972; McGill, 1963; Restle, 1961) have been concerned with the distribution of RT within an individual over a series of trials. McGill (1963) has assumed that (*a*) stimulus intensity is a random variable, (*b*) the rate of neural impulses associated with a stimulus increases as stimulus intensity increases, and (*c*) the subject's response criterion can be expressed as a criterion number of impulses. On this basis, an inverse relationship between RT and intensity may be derived. Grice (1968) has questioned the validity of McGill's model as proposed, since it requires that the subject's response criterion remain constant over a large block of trials. Instead, he has offered an alternative model in which the criterion, but not the signal intensity, is assumed to vary from moment to moment or trial to trial. Both models make predictions about the RT-density function as well as the mean of the distribution. Both assume random, normal distributions.

Luce and Green (1972) have also provided a mathematical theory which attempts to predict sensory decisions and their associated latencies. The underlying density function is assumed to be Poisson, although that is not a critical requirement of the approach. These authors compared several possible models for each of a number of decision rules against a variety of experimental situations. In general, the experimental conditions considered impose requirements which are different from the conventional RT situation, for example, the requirement for a discrete secondary stimulus which acts as a marker to indicate termination of the primary stimulus.

METHOD

An attempt was made to find and evaluate all studies published in English which used RT as a dependent measure. A few unusually important papers in French and German were also included. The restrictions then imposed on the papers before they could be accepted for further use were the following:

1. Viewing was restricted to central vision. Studies which did not specify either monocular or binocular viewing were assumed to have been binocular.

2. The spatial position of the signal had to be constant and known by the subject.

3. The time of arrival of the signal had to be known by the subject within close tolerance. Studies using both random and nonrandom foreperiods were accepted if it could be assumed for the former that the subjects were well practiced.

4. There could be only one possible signal and one possible response.

5. The report had to provide either the actual values of luminance, duration, visual angle, and background luminance, or information with which those values could be derived.

6. The procedures and experimental design had to provide unconfounded data.

On this basis, a total of 14 studies were accepted as providing usable data. Those data were converted to common units of measurement of the stimulus parameters and of RT, namely, luminance in millilamberts, size or "area" of stimulus in degrees of visual angle, and stimulus duration and RT in seconds. All data are reported in those units.

RESULTS AND DISCUSSION

The first step was to plot the data of all studies together as a function of stimulus luminance in the hope that the effects of other independent variables would be apparent. On this basis, it seemed that luminance, duration, and area interacted in a very complex way, and that the variable of greatest effect was probably luminance. A second step was to manipulate the dependent variable in the hope of separating phenomena which might be components of the total RT. In particular, attention was given to the question of processes which might be going on during the presence of the stimulus and those which might follow its offset. One approach to this distinction was simply to subtract the stimulus duration (t) from RT and to compare RT with RT $- t$. If RT does not depend on t, or if Bloch's Law holds for latencies measured from the termination of the stimulus (Grossberg, 1968), then a comparison of these two dependent measures might yield different functions. However, they yielded the same trends except at relatively long durations and large luminances where RT was essentially constant and, therefore, where RT $- t$ dropped rapidly toward very small values. This suggested only that RT does decrease with increasing t.

Another dependent measure which was considered was the fraction, t/RT, which represents the proportion of the total RT in which the stimulus was present. The quantity, $1 - t$/RT, then, represents the proportion of the total RT during which the stimulus is absent. Use of these ratios tended to organize the data very systematically, but ultimately it was decided that they provided no more information than was available in the conventional RT measure. Therefore, they were not pursued further. The data reported are all in terms of the conventional measure.

Effects of Area

Figure 1 compares the effects of area for different, but comparable, values of duration as a function of luminance. The foveal-sized data of Grossberg (1968) cover the lower luminance range; the data of three other studies using larger stimuli represent higher luminances. In both cases, RT is a decreasing function of luminance. It may also be seen that duration is generally an effective variable, and that the larger areas were associated with shorter RTs.

Although Figure 1 suggests that RT is inversely related to size of stimulus, not enough data are shown to suggest the nature of the area function. This is characteristic of the foveal data which we were using. That is, with the exception of a single data point from Poffenberger (1912), who used one degree, all of the studies used employed stimuli which were well within foveal dimensions, .5 degrees of arc or less, or which exceeded foveal dimensions within the range of 1.1–4.0 degrees of arc. Not enough variation within these limits was available for developing an area trend. For that reason, some of the analysis to follow distinguish only between foveal-sized (≤ 1.0 degree) and extrafoveal-sized (> 1.0 degree) stimuli. That is, the data are treated within each of those categories as if area has no effect on RT. We conclude that there is an effect between categories, but that not enough data

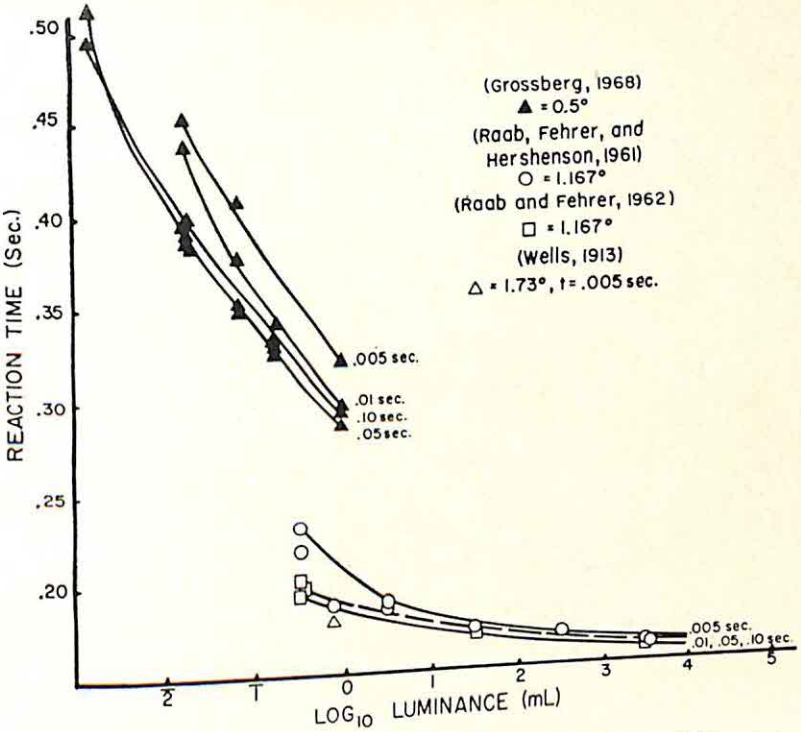


FIG. 1. Effects of signal luminance, duration, and size on RT. (Viewing was binocular. Lines were drawn by eye.)

are available to determine what effects, if any, exist within these two categories.

Luminance, Duration, and Onset versus Termination

The durations used varied from reported $t = .00001$ second to durations which nominally exceeded RT. An attempt was made to view the luminance function in an uncomplicated way by grouping the durations used into class intervals and treating all data within an interval as the same in duration. An example is shown in Figure 2 using the interval .002–.010 second. The figure also provides a comparison of the use of stimulus onset with termination as the signal for the data of Sticht (1969). Regarding that comparison, stimulus onset is shown to be associated with a faster response. This was not a consistent result, however, even for the other data of Sticht, not shown here. In any case, the termination data are easily included within the onset data, which suggests that if there is a difference, it represents the effect of a minor variable. As a precaution, however, the data reported below are only for the case in which stimulus onset was used as the signal.

As in the previous figures, Figure 2 shows that extrafoveal target sizes produced shorter RTs. There is an indication again that short durations tend to be associated with longer RTs. As before, RT tends to decrease with luminance. This is certainly the case within studies. Between studies, duration and size appear to interact. A variety of other plots using different duration intervals provided essentially the same trends except, as noted, they were inconsistent about the effects of onset and termination.

Although plots of the type shown in Figure 2 were informative, they were considered to be too confounded by the effects of area and not really clear about the effects of duration. All of the data, therefore, were plotted as a function of duration with luminance as the parameter. This was done separately according to the following conditions: foveal size monocular, foveal size binocular, and

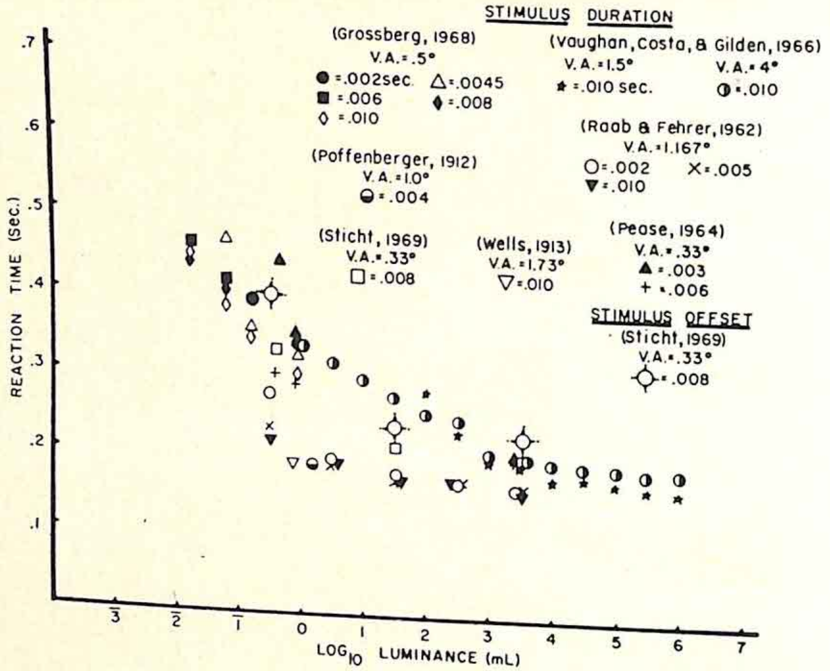


FIG. 2. RT as a function of luminance for varying signal sizes and durations varying within the range .002–.010 second. (Data from both monocular and binocular viewing conditions are included. All data are from response to stimulus onset except the three indicated points from Sticht, 1969.)

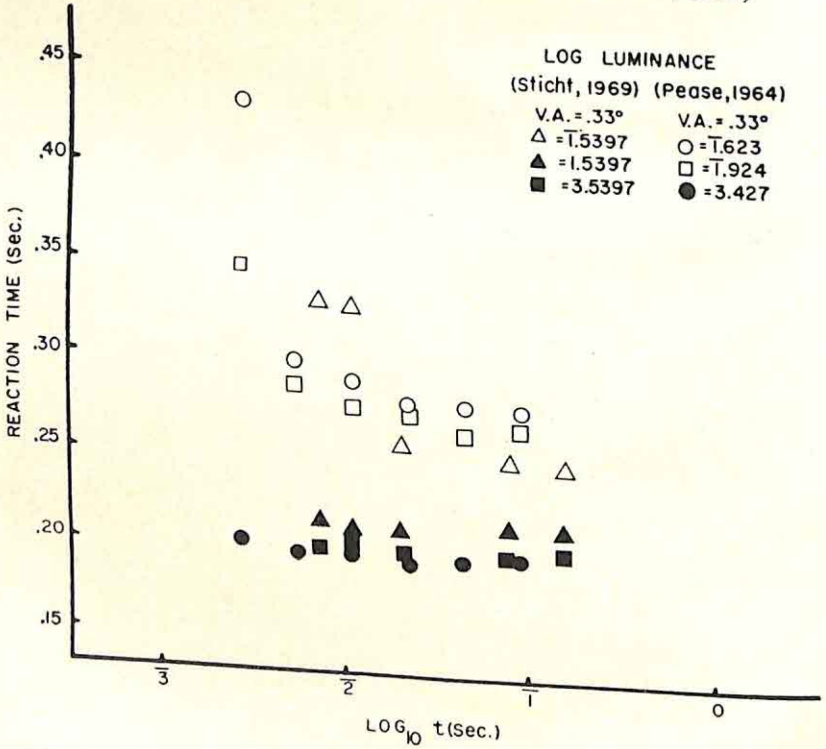


FIG. 3. RT as a function of stimulus duration and luminance. (Viewing was monocular; signal size was .33 degrees. Data from Sticht, 1969, and Pease, 1964.)

extra foveal size binocular. The results are shown in Figures 3, 4, and 5, respectively. No monocular studies using signals of extra foveal size were available.

Comparison of Figures 3, 4, and 5 indicates that RT is a negatively accelerated decreasing function of stimulus duration, and that RT decreases as luminance increases. From Figures 1 and 2 it was already seen that the luminance function is negatively accelerated and decreasing. It is also apparent that the separation of the data into the three figures according to size and monocular versus binocular viewing resulted in a considerable gain in ordering of the data. Since the data fell most systematically when plotted as time functions with luminance as a parameter, it was decided to attempt a formal expression for the data in that form.

When the data of Figures 3 and 4 were plotted on one graph, it was found for approximately comparable luminances and sizes that the monocular RTs were shorter in general than the binocular ones. This violates any reasonable expectation. The monocular RT should be longer than or equal to the binocular RT; it should not be shorter. Nothing in the published reports suggests that one or the other data sets should be considered less reliable. The decision that the monocular data were probably less reliable was made on the basis of a comparison of inversions within the two figures and on the greater comparability of the time and luminance trends of Figure 4 to those of Figure 5, which represents binocular data for the larger areas. The differences btween the monocular and binocular data are reconsidered below in a different context.

Inspection of Figure 4 at the two highest luminances suggests that perhaps two time functions may be required, one for very short durations and one for longer ones. Although that possibility is not consistent within the figure, an attempt was made to treat the data at each luminance as a two-factor system. This not only could not be done consistently for the family of relationships in the figure, but even for the two highest luminances it could not be done well with consistent functional forms. The possibility was abandoned, therefore.

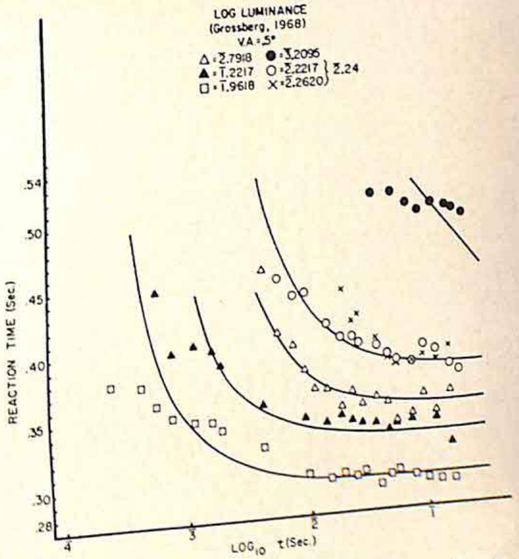


FIG. 4. RT as a function of stimulus duration and luminance. (Viewing was binocular; signal size was .5 degree. Data from Grossberg, 1968. Curves calculated with Equations 2 and 3.)

Of the various functions which might have described the data of Figures 4 and 5, the expression found to be most consistently adaptative within and between the figures was

$$\mathrm{RT} = \frac{t}{bt - a}. \qquad [2]$$

In this expression the constant, a, depends on the value of duration at which RT approaches infinity (t_∞), and the constant, b, is a function of the asymptotic value of RT. The latter can be seen in the figures as depending on luminance. The first steps, therefore, were to estimate both kinds of asymptotes for the data of Figure 4. This was done visually except for t_∞ for the lowest luminance where t_∞ was determined from $t_\infty = -a/b$. The constant, a, for that luminance was determined iteratively.

Obtaining b required first that the luminance function be developed. Asymptotic values were estimated for each curve of Figure 4 with the exceptions of $\overline{2}.22$ and $\overline{2}.26$, which instead were treated together as $\overline{2}.24$. The estimated asymptotic values were then plotted as a function of luminance in various

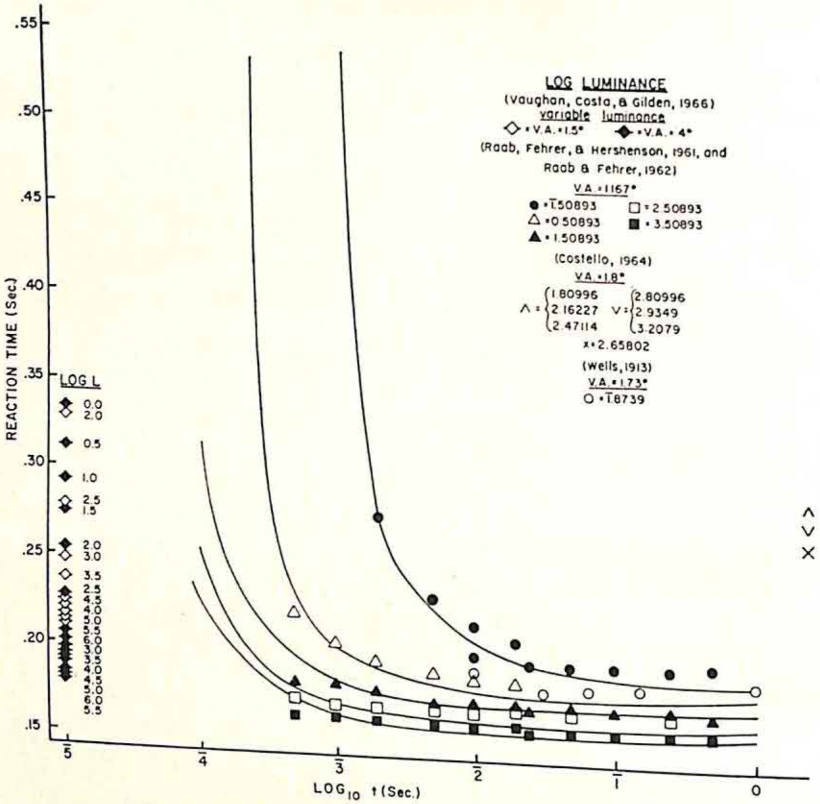


FIG. 5. RT as a function of stimulus duration and luminance for extra-foveal-sized signals viewed binocularly. (Curves calculated with Equations 2 and 4.)

transforms. The best fit was clearly a power function. That relationship in the form of the reduction line is shown in Figure 6. The equation was fitted graphically. Its antilogarithm is

$$K = .302L^{-.06}, \qquad [3]$$

where K is the asymptotic value of RT in seconds at a given luminance, L is luminance in millilamberts, and .302 is a constant. No attempt was made to add a constant to Equation 3 which would impose a limit on it.

The value of b at any luminance was obtained for Equation 2 by solving Equation 3 for K at that luminance and using the relationship, $b = 1/K$.

Figure 7 presents the values of t_∞ used as a function of luminance. The smooth line,

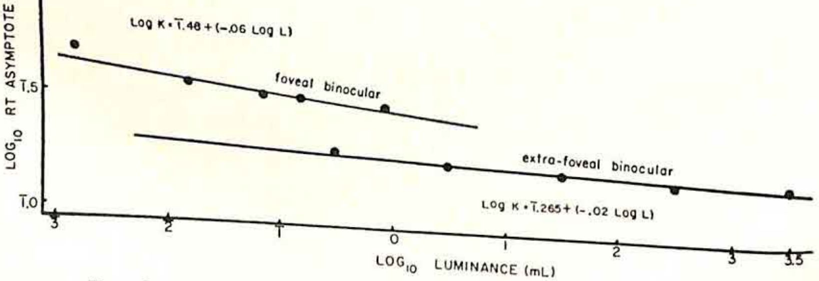


FIG. 6. Log-log plot of asymptotic RT versus luminance for binocular viewing.

drawn by eye, suggests a power function for this signal size. Using Equations 3 and 4 and the relationship $-a = bt_\infty$, the constants required for Equation 2 may be obtained for any luminance. When this was done for the data of Figure 4, the curves obtained tended to overestimate slightly, so a subtractive constant of fit was introduced. With this constant, Equation 2 as applied to Figure 4 becomes

$$\mathrm{RT} = \frac{t}{bt - a} - .01. \qquad [2a]$$

It may be seen in Figure 4 that the fit is poorest at the lowest luminance. Whether it would improve were more data available there cannot be stated. At the two highest luminances, the fit misses the points representing the shortest durations. This does not seem to be too serious since those points are close to t_∞ and, therefore, represent very low visibility levels. They may be expected, therefore, to be the least reliable of the points and, in fact, this is suggested, since they can be viewed as not in line with their own trends.

Using the methods described above, Equation 2 was also fitted to the data of Figure 5. The equation for the antilogarithm of the asymptotic luminance function shown in Figure 6 is

$$K = .1799L^{-.02}. \qquad [4]$$

Shown in Figure 7 are the values of t_∞ obtained. The smooth line was drawn through the points by eye. It suggests that a different function describes the dependence of t_∞ on luminance for the larger stimulus sizes than for the smaller ones. The difference may be the result of the fitting process, since it was possible to estimate t_∞ graphically for more curves at the smaller sizes than for the large ones. On the other hand, the difference in functions could be a genuine difference between within-foveal and extrafoveal viewing conditions. In applying Equation 2 to Figure 5, no fitting constant was used.

The curves shown in Figure 5 provide a reasonable first approximation of the data reported over the range of stimulus durations, $\mathrm{RT} > t \overline{>} .0005$ second. The failure to describe the data of Costello (1964) is to be expected, since there is no way to estimate a value for the abscissa except that it should be no greater than the RTs reported and, undoubtedly, less. The microsecond duration data were also not accounted for by Equation 2.

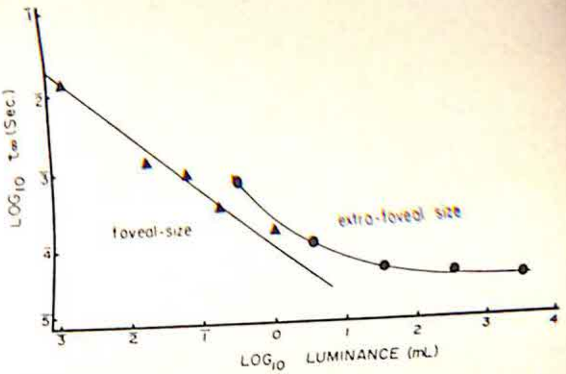


FIG. 7. Log-log plot of t_∞ versus luminance for binocular viewing.

Attempts to fit other functions to the data of Figures 4 and 5 were less successful than Equation 2. As noted, the data of Figure 3 for monocular viewing were not sufficiently consistent to allow for the fitting of a single function. The inconsistency lay not so much in the shape of the duration functions shown in the figure, but in the function to be used to predict the asymptotic RT, that is, K; the luminance trend obtained did not make sense when compared to the relationships found for the other two figures.

Energy Relationships

If RT is a constant value for a constant stimulus energy, then Bloch's Law holds for the visual RT. If so, at least within the limits of application of the law, RT may be viewed as largely, if not entirely, dependent on the same retinal process as that on which visual sensitivity depends. If the law does apply, Figures 4 and 5 suggest critical durations which depend on luminance. On the other hand, if RT does not behave as a visual threshold by following Bloch's Law, questions about the relative importance of peripheral and central processes are appropriate.

In the visual literature, Bloch's Law is tested by plotting the product of luminance and duration at threshold as a function of duration. The constant dependent measure in that case is the threshold which is defined independently of both luminance and duration. A comparable procedure cannot be employed in evaluating the law for RT, since the dependent measure, RT, includes duration as a component value. That is, for this purpose the independent and the dependent variables are completely confounded so that the duration quantity would appear on the abscissa, in the product on the ordinate, and in the referent constant RT.

To circumvent the problem, RT investigators have attempted to make use of the finding that the critical duration up to which a constant energy is found at threshold varies with other parameters of the visual stimulus. Since the critical duration of the RT curve varies with luminance, as shown in Figures 3, 4, and 5, investigators have used that phenomenon to support the applicability of Bloch's Law. This of course is a *non-sequitur*. The test may be made, however, simply by plotting RT as a function of the product of luminance and duration and observing whether or not RT is a constant for constant abscissa values of the graph. This has been done in Figure 8 using all of the data of Figures 3, 4, and 5 except those reported with $t > \text{RT}$. It is clear that Bloch's Law does not hold. The same conclusion was reached by Grossberg (1968), whose data are included. Figure 8, which shows that the conclusion is general, also indicates that as *It* increases, RT decreases in a negatively accelerated manner.

There is also some suggestion in Figure 8 that at longer RT values there may be a family of curves whose asymptotes vary with luminance, while for the shorter RTs another family of curves might possibly depend on area. The smooth lines, drawn by eye, are intended to represent the minimum RT for the binocular and the monocular data. A comparison of those curves suggests that the monocular condition did, in fact, tend to produce longer RTs. The smooth curve for the binocular condition, then, defines an envelope within which all RTs are likely to fall when *It* is the independent variable. The monocular curve may be interpreted similarly. These lines may be used to estimate the shortest RT to be expected at any *It* regardless of other stimulus or viewing conditions.

Whereas the product of duration and luminance represents the energy per unit area, the product of luminance, duration, and area represents the total energy. A plot of RT versus *LtA* was so similar to that of Figure 8 that it is not shown. Of additional interest in that plot, as well as in Figure 8, was that the RTs obtained at microsecond durations were much more in line with the remaining data than they are in Figure 5.

Definition of the Stimulus

So far we have presented an analysis of RT based on a definition of the stimulus as an energy applied for a period of time, and of the response as an event which occurs when a necessary amount of energy has been applied. RT is defined as the latency of the response measured from the onset of the stimulus. A critical underlying assumption appears to be that when the external stimulus energy is removed, stimulation stops and that the temporal lag observed before onset of the response is due to lags in the nervous and muscular systems. With those assumptions it is clear that a reasonably consistent, quantitative literature has developed, and that it can be organized within limits into simple mathematical relationships.

On the other hand, the assumption that stimulation ends as defined by its externally controlled duration is unacceptable. A distinction between the distal and proximal stimulus seems particularly pertinent here, as well as in studies of visual detection, because it is apparent in Figure 8 that micosecond stimulus durations produce RTs which are reasonably located within the range of effect of other stimulus durations. To achieve such results requires a persistence of the visual input after the external energy is removed. Under ordinary stimulating conditions, this persistence may last for at least .25 second (Mackworth, 1963; Sperling, 1960), and under conditions which produce afterimages, it might last for several seconds. During this period, informa-

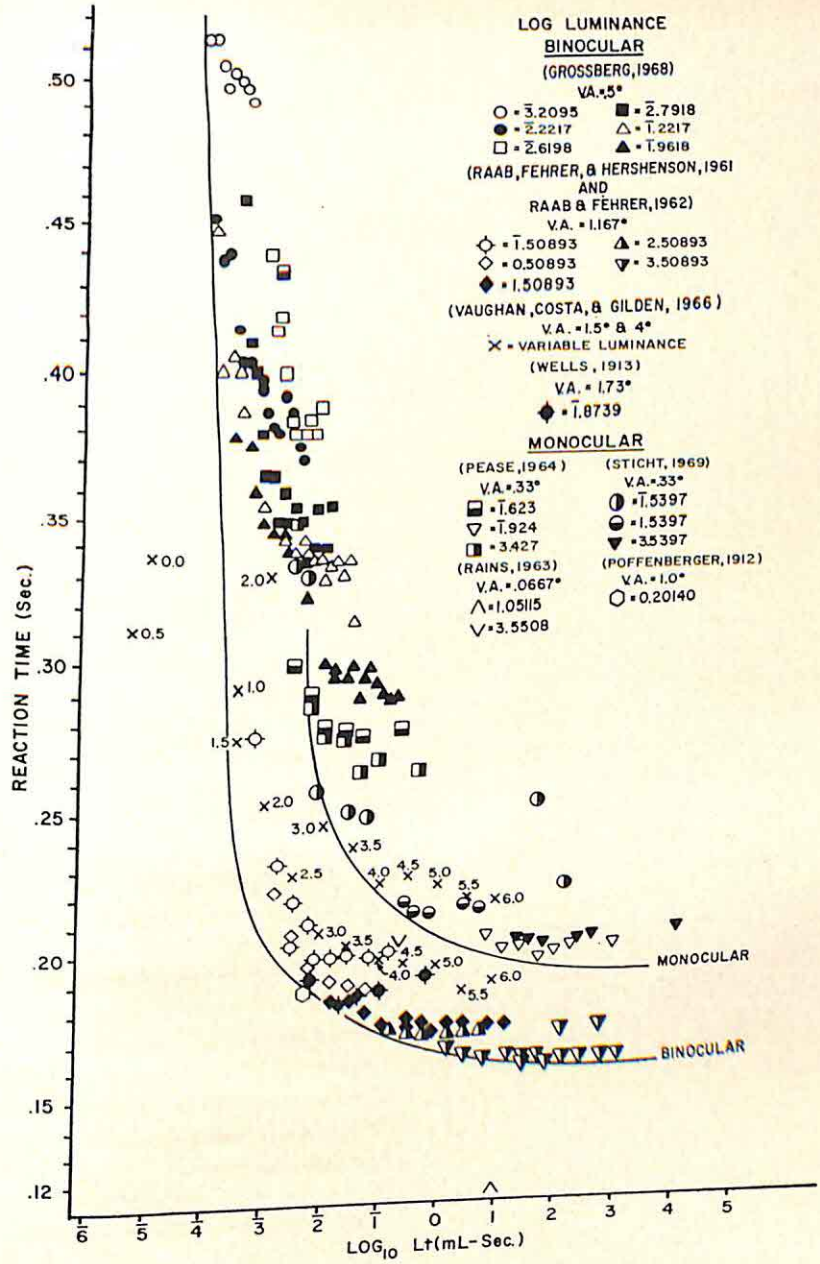


FIG. 8. RT as a function of log stimulus energy.

tion can be extracted by the individual (Teichner & Wagner, 1964).

Response Criterion Model

Figure 9 presents the fundamental relationships which the theories of Grice (1968) and McGill (1963) appear to have assumed in common. We shall not be concerned with the distribution functions of those models, but rather we shall assume that if either the input or the criterion (or both) is not constant from moment to moment or trial to trial, then their averages are, and it is the average values with which we shall deal. On this basis, Figure 9 assumes that the individual's response criterion is represented by an accumulated number of neural impulses so that the greater the luminance,

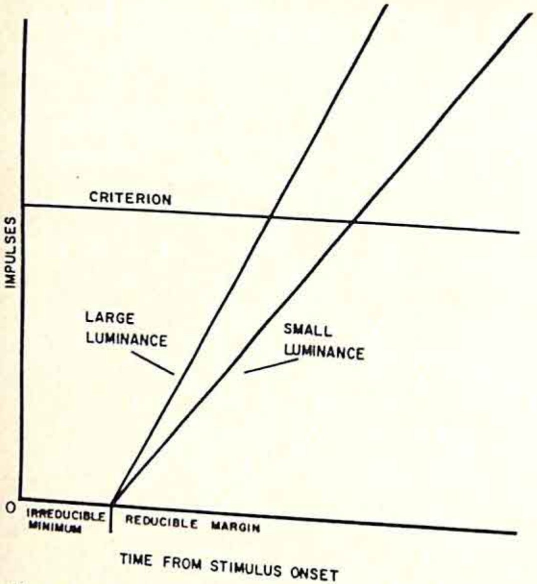


FIG. 9. Graphical representation of Grice's RT model without distribution functions. (Adapted from Grice, 1968.)

and therefore the greater the pulsing rate, the sooner the criterion is reached.

The abscissa of Figure 9 represents RT, that is an irreducible minimum plus the duration of pulsing required to reach the response criterion (RC). Designating the reducible portion as $\mathrm{RT}' = \mathrm{RT}$ minimum, the model states that pulsing rate (PR) and luminance (L) are directly related. It is evident from the model that

$$\mathrm{RC} = \mathrm{PR} \times \mathrm{RT}' \qquad [5]$$

therefore, if

$$\mathrm{PR} = kL \qquad [6]$$

then

$$\mathrm{RC} = kL \times \mathrm{RT}' \qquad [7]$$

and the product of L and RT' is constant for all values of L.

Equation 7 can be evaluated empirically by inspection of Figure 10 which relates $\mathrm{RT} \times L$ to duration of the distal stimulus with luminance as a parameter.[3] The data of all of the studies used for Figure 8 are incorporated in this figure. The lines were drawn by eye. It is apparent from the figure that $\mathrm{RT} \times L$ is not constant and, therefore, that the linear assumption of Equation 6 is not adequate.

Inspection of Figure 10 shows that the vertical arrangement of the lines is due almost entirely to luminance. Looking over this arrangement, it may be seen that all of the data are systematically ordered by luminance, with minor exceptions. Previously noted differences between signal size and monocular versus binocular viewing are also minor except at the shortest duration. Although the large range of $\mathrm{RT} \times L$ reflects good experimental coverage of the luminance variable, it is unfortunate that more comparisons are not available where the luminance differences are small. However, for those few cases where the differences are small, especially for those smaller than the RT range, luminance is still the major factor.

Figure 10 also shows that the effect of duration of the distal stimulus was to decrease $\mathrm{RT} \times L$ as duration increased. Data which could not be handled in Figure 5 may be seen here to fall on single trends. The data of Costello (1964) are incorporated in this context, since now it is RT which defines the effective stimulus duration, so that even though those data are at large values on the abscissa, they are meaningfully placed in terms of the ordinate. They may be seen simply to be at the asymptote of the duration function. All in all, the figure shows that duration affected $\mathrm{RT} \times L$ systematically, though not with a large effect, at low to moderate luminances, and little or not at all at higher luminances. In the sense that $\mathrm{RT} \times L$ represents the response criterion, the figure shows that the response criterion varies widely with luminance and to a small degree with exposure time.

Figure 11 provides a plot of $\mathrm{RT} \times L$ as a function of luminance. The points were obtained from Figure 10 as single-point estimates of the asymptotes of the duration function where such functions were available. Where only one point was available to represent a given luminance or size, that point was used as if it were the asymptotic value. Where two points were available, an aver-

---

[3] Theoretically the ordinate of Figure 10 is $(\mathrm{RT}' \times L) +$ minimum.

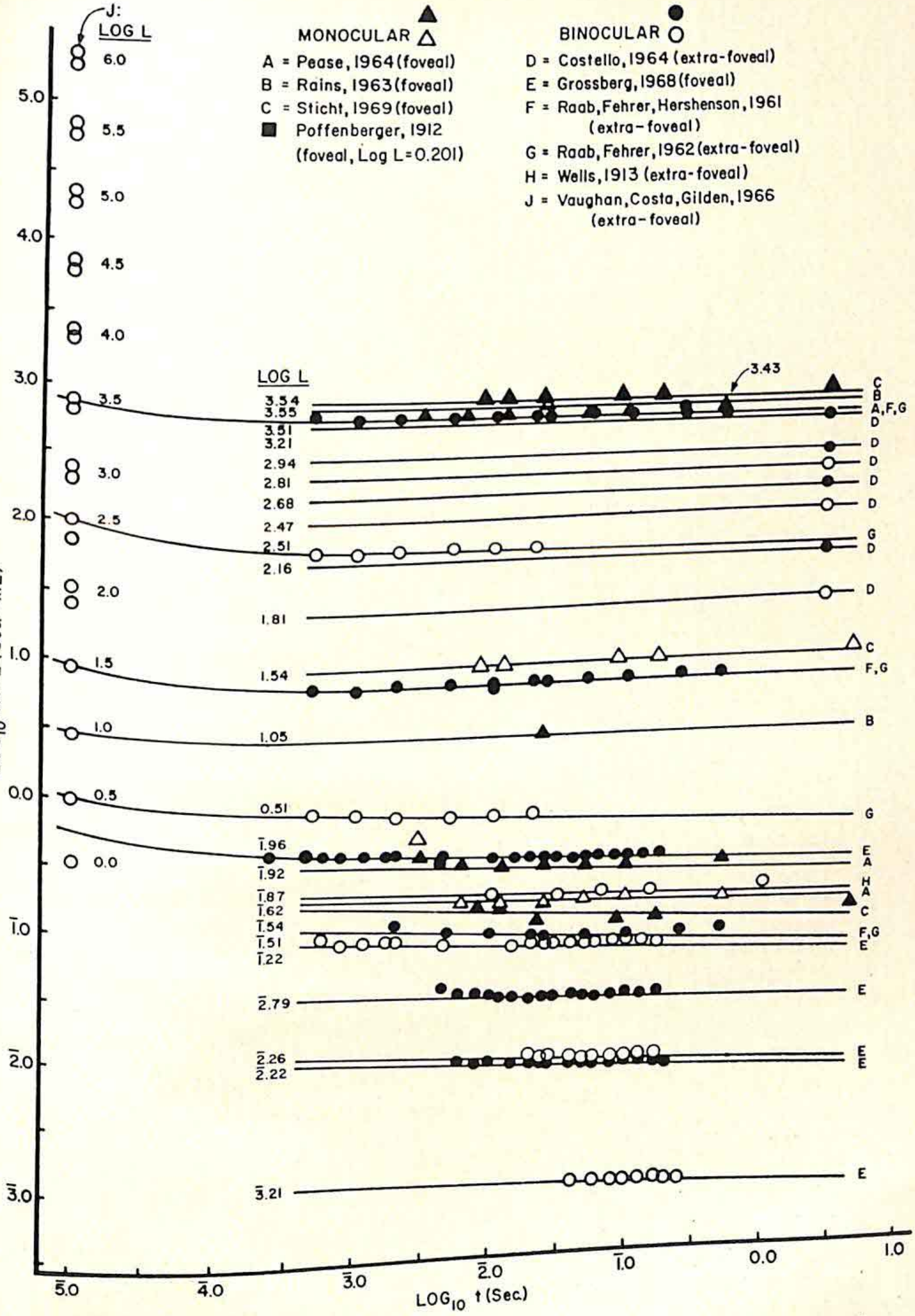


FIG. 10. Response criterion expressed as log (RT × L) as a function of log stimulus duration and log luminance. (Smooth lines drawn by eye. For all pairs of points at log $t = \bar{5}.0$, the upper points are based on a stimulus of 1.5-degree visual angle, the lower for a stimulus of 4.0 degree: the single points at this duration for luminances below log 2.0 millilamberts are based on the 4.0-degree stimulus.)

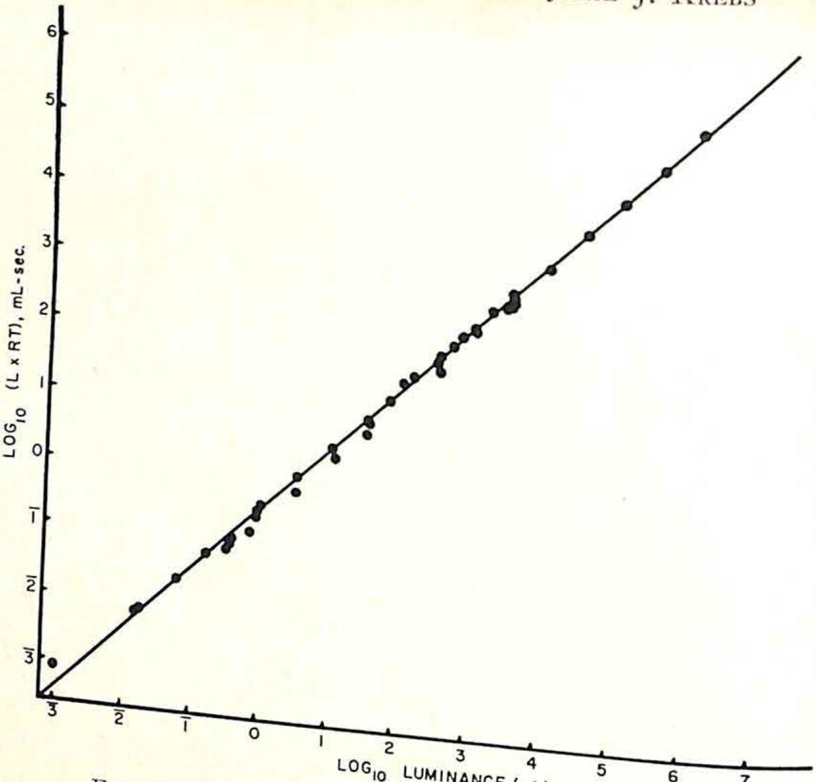


FIG. 11. Log (RT × L) as a function of log luminance.

age value was used. The figure shows that log RT × L is proportional to log luminance. The antilogarithm of the equation for the line shown is

$$\mathrm{RT} \times L = .302L^{.96}, \qquad [8]$$

which indicates again that RT × L is almost entirely determined by luminance. Dividing by L in Equation 7,

$$\mathrm{RT} = .302L^{-.04}, \qquad [9]$$

for which the exponent is midway between that of Equations 3 and 4, and the constant is identical to that of Equation 3. Thus, for empirical prediction, Equation 9 provides a general luminance function for RT which can be used with Equation 2 if factors other than luminance are ignored.

Equations 8 and 9 also provide a basis for developing more adequate assumptions for the response criterion model than those expressed above. Equation 8 indicates that the response criterion increases as a power function of luminance. If pulsing rate is also a power function of luminance, then the resulting RT will be dependent on the difference between the functions. That is, as luminance increases, pulsing rate increases, and this will operate to reduce RT. At the same time as luminance increases, the response criterion will increase, and this will operate to lengthen RT. For RT to become shorter as luminance increases, pulsing rate must increase faster than the criterion. In general,

$$\mathrm{PR} = kL^{m} \qquad [10]$$

and

$$\mathrm{RC} = aL^{n} \quad m > n \qquad [11]$$

which yields

$$\mathrm{RT}' = cL^{n-m} \quad c = a/k, \qquad [12]$$

the constant and exponent for which is provided by Equation 9.[4]

---

[4] It should be noted that any values of m and n which differ by .04 will provide Equation 9 and will allow the log-log plot of Equation 8 to have a slope of .96.

If the criterion event is to be expressed in energy terms, it should probably be expressed as a function of change in energy or, if in neurophysiological terms, as a change in pulsing rate rather than as a cumulated number of pulses. This follows from the finding that there is no difference in RT for response to the onset or to the termination of a flash of light. Even if there were a difference, the cumulated pulse concept would not indicate how a response could be made to stimulus termination without postulation of a second hypothetical process (e.g., "off" receptors). We note then that the neurophysiological criterion event should be a criterion acceleration or deceleration of a process. For the present, however, the cumulated impulse model provides a fruitful behavioral approach.

Within the framework of the response criterion model, the results appear to have implications for the view that RT is the sum of a sensory lag, an energy or intensity processing lag, and a lag which depends on attentional, cognitive, and motivational factors. The finding that the criterion varies with luminance and with duration, both of which are energy factors, suggests that two kinds of energy-processing activities are involved. One of these must be a criterion-development activity, and one a "count-down" of impulses. Since the counting activity must begin as soon as the impulses are available to it (after some minimal stimulus time), and since the criterion-development activity should also be initiated at that time, these two processing activities must go on in parallel. Furthermore, any other factors important to the development of the criterion must be effective during the same period. Thus, using this model and the present results, RT seems to depend on a serial–parallel system, that is, one in which parallel processing activities follow a sensory lag.

REFERENCES

Bartlett, N. R., & McLeod, S. Effects of flash and field luminance upon human reaction time. *Journal of the Optical Society of America,* 1954, **44**, 306–311.

Berger, G. O. Ueber den Dinfluss der Reizstärke auf die Dauer ein facher psychiser Vorgänge mit besonder Rücksicht auf Lichtreize *Philosophische Studien,* 1886, **3**, 38–93.

Cattell, J. M. The influence of the intensity of the stimulus on the length of the reaction time. *Brain,* 1886, **8**, 512–515.

Costello, C. G. The effects of depressant and stimulant drugs on the relationship between reaction time and stimulus light intensity. *British Journal of Social and Clinical Psychology,* 1964, **3**, 1–5.

Grice, G. R. Stimulus intensity and response evocation. *Psychological Review,* 1968, **75**, 359–373.

Grossberg, M. The latency of response in relation to Bloch's Law at threshold. *Perception and Psychophysics,* 1968, **4**, 229–232.

Hull, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychological Review,* 1949, **56**, 67–76.

Luce, R. D., & Green, D. M. A neural timing theory for response times and the psychophysics of intensity. *Psychological Review,* 1972, **79**, 14–57.

Mackworth, J. F. The duration of the visual image. *Canadian Journal of Psychology,* 1963, **17**, 62–81.

McGill, W. J. Stochastic latency mechanism. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963.

Pease, V. The intensity-time relation of a stimulus in simple visual reaction time. *Psychological Record,* 1964, **14**, 157–164.

Piéron, H. Nouvelles recherches sur l'analyse du temps de latence sensorielle et sur la loi qui relie ce temps a l'intensite de l'excitation. *Année Psychologique,* 1920, **22**, 58–142.

Poffenberger, A. T. Reaction time to retinal stimulation with special reference to the time lost in conduction through nerve centers. *Archives of Psychology,* 1912, **23**, 1–73.

Raab, D., & Fehrer, E. The effect of stimulus duration and luminance on visual reaction time. *Journal of Experimental Psychology,* 1962, **64**, 326–327.

Raab, D., Fehrer, E., & Hershenson, M. Visual reaction time and the Broca-Sulzer phenomenon. *Journal of Experimental Psychology,* 1961, **61**, 193–199.

Rains, J. D. Signal luminance and position effects in human reaction time. *Vision Research,* 1963, **3**, 239–251.

Restle, F. *Psychology of judgment and choice.* New York: Wiley, 1961.

Sperling, G. A. The information available in brief visual presentations. *Psychological Monographs,* 1960, **74**(11,Whole No. 498).

Sticht, T. G. Effects of intensity and duration on the latency of response to brief light and dark stimuli. *Journal of Experimental Psychology,* 1969, **80**, 419–422.

TEICHNER, W. H. Recent studies of simple reaction time. *Psychological Bulletin,* 1954, **51**, 128–149.

TEICHNER, W. H., & WAGNER, M. Visual afterimage as a source of information. *Human Factors,* 1964, **6**, 141–156.

VAUGHAN, H. G., JR., COSTA, L. D., & GILDEN, L. The functional relation of visual evoked response and reaction time to stimulus intensity. *Vision Research,* 1966, **6**, 654–656.

WELFORD, A. T. The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics,* 1960, **3**, 189–230.

WELLS, G. R. The influence of stimulus duration on reaction time. *Psychological Monographs,* 1913, **15**(5, Whole No. 66).

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology.* New York: Holt, 1954.

(Received November 29, 1971)

Psychological Review
1972, Vol. 79, No. 4, 359–367

SUBJECTIVE CONTOURS AND APPARENT DEPTH[1]

STANLEY COREN[2]

New School for Social Research

The phenomenal appearance of contours in the absence of abrupt stimulus gradients was first discovered by F. Schumann. Recently such contours have been produced in Julesz patterns using binocular disparity. Analysis indicates that both monocular and binocular subjective contours result from the presence of depth cues in the stimulus array.

Schumann (1904) investigated the physical parameters which affect the perception of contour. He reported that for a contour to be perceived the prerequisite was some relatively abrupt local change in brightness or color. In the course of these investigations, however, he noted that under certain favorable conditions contours could be made to appear in areas of the visual field where the physical stimulation was homogenous. As evidence of this, he presented the pattern shown in Figure 1. In the center of this pattern most observers see a white square, bounded by intermittent black and white on either side. Since the contours delineating this square are not actually present in the stimulation, these phenomena have been called "subjective contours" (Osgood, 1951).

The subjective contours in the Schumann figure are somewhat unstable. Not all observers see these contour lines, and when they are carefully fixated, they tend to disappear. Their appearance is generally enhanced when the configuration is of small visual angle. All of these factors led Schumann to conclude that subjective contours were the result of some organizational process, similar in nature to closure, which resulted in the perception of a square bounded by contours which were not physically present.

Schumann's work on subjective contours has been extended by Kanizsa (1955) who created a number of configurations which produce much more compelling contours than those in Figure 1. In addition, Kanizsa was able to demonstrate that curvilinear as well as straight contours could be produced. Some representative examples of patterns which produce subjective contours are shown in Figure 2.

When observers look at Figure 2A, which has been adapted from Kanizsa (1955), they invariably report seeing a white triangle in the center of the configuration, although no such figure physically exists. The properties of this triangle are rather interesting. In addition to being whiter than the surround, *it is almost always described as a plane surface which appears to be in front of the other elements in the pattern.* The triangle is apparently opaque and capable of obstructing the observer's view of the pat-

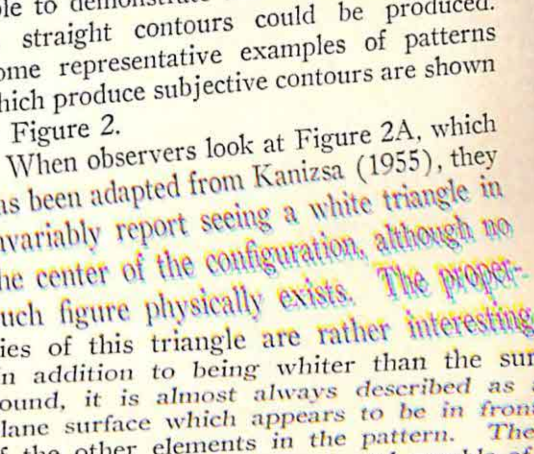


FIG. 1. The Schumann figure: A central white square is seen in this configuration, especially when the figure is of small visual angle.

---

[1] This work was partially supported by National Institutes of Health Grant No. 16327 which was awarded to Leon Festinger. I would like to thank Leon Festinger for his invaluable criticism of this work. A number of people have commented on this work at various stages in its development; these include Joan S. Girgus, Lloyd Kaufman, Charles S. Harris, Julian Hochberg, and the members of the Program in Visual Perception. Saulo Sirigatti assisted in translation of some source material.

[2] Requests for reprints should be sent to Stanley Coren, 66 West 12th Street, New York, New York 10011.

tern elements behind it. Figure 2B shows a variant in which a subjectively produced white rectangle appears in the center, while Figure 2C shows that curvilinear subjective contours are also possible. Figure 2D is the negative form of Figure 2A. It is clear that with the brightness of the lines and background reversed, the subjective triangle is still seen; only here note that it appears as blacker than the black of the background. The intriguing aspect of these figures is that large stretches of the perceived contours are not physically present, but rather are curved out of homogeneous portions of the stimulus field by the perceptual system.

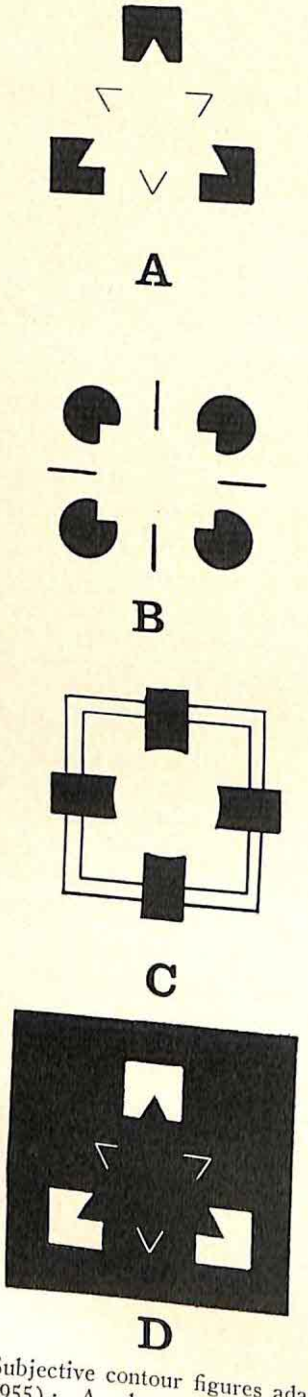


FIG. 2. Subjective contour figures adapted from Kanizsa (1955): A shows a subjective white triangle; B, a square; C illustrates that curvilinear subjective contours may be produced; and D shows a subjective contour with reversed brightness of that in A.

Virtually all observers see the subjective contours in these configurations. Knowledge of the fact that there is no triangle physically present in Figure 2A does not seem to lessen the strength of the percept, and the contour remains visible even with careful fixation. As in the case of Figure 1, figures of small visual angle tend to produce an enhanced effect.

The phenomenon of apparent contour has never been convincingly explained. Kanizsa (1955) reiterates a Gestalt position and refers to subjective contours as unusually strong examples of closure or figure-ground organization. His position is close to Schumann's. This same theoretical stand was also taken by Rubin (1921).

Binocular Disparity and Subjective Contour

Julesz (1964) presented a stimulus array which produced a very compelling subjective contour in the absence of an abrupt change in brightness. Using computer matrices of randomly generated dots, he created a sensory environment devoid of all depth and familiarity cues except retinal disparity. Stereoscopic presentation of identical dot matrices results in the perception of a simple two-dimensional display of dots on a homogeneous background. If the dots are small enough or slightly blurred, the perception is of a homogeneous textured field. If, however, these displays are altered by shifting the location of all dots in a central submatrix one column laterally, the percept immediately changes. If, for instance, the dots in a square central area seen by the left eye are shifted one column to the right, the observer now reports that this central square region is seen in depth. It is now localized as being closer to the observer than the background. In addition to this depth effect, the observer notes that this region is bounded by subjective contours which appear to be localized between the rows and columns of dots. As White (1962) pointed out, these contours are clearly subjective in nature and do not require the usual stimulus correlates associated with the perception of contours. Closing either eye results in an immediate loss of the impression of depth and a disappearance of the contour.

Under optimal conditions, these contours are very impressive and are indistinguishable from those produced by sharp stimulus gradients. Shipley (1965) has shown that even in blurred or out-of-focus configurations, the sharp subjective contour still accompanies the perception of the central form in depth.

Lawson and Gulick (1967) extended the analysis of subjective contours seen in these patterns. They presented stimulus arrays in which the disparity cue produced subjective edges bounding large homogeneous areas of the field. They have shown that the important feature of the stimulus display is not simply retinal disparity, for in some cases they were able to produce arrays which resulted in the perception of a group of dots seen in depth, not bounded by any contours, but rather form disparity. Form disparity is defined here as that pair of disparate views that would result if an opaque form were physically present in the visual field and were differentially obstructing parts of the background matrix which were seen by the two eyes. A striking example of this is found in the work of Kaufman (1964) and Kaufman and Pitblado (1965), who were able to show that even when the elements presented to the two eyes were quite different in nature (i.e., were rows and columns of different letters), the effect of disparity was the perception of a target in depth, bounded by subjectively seen contours. Lawson and Gulick (1967) observed that in the absence of the perception of depth they found no subjective contours.

The question which immediately springs to mind is whether the subjective contours visible in these stereograms are due to the operation of the same general mechanism as are the monocularly visible contours in Figure 2. One hint that such might be the case comes from the phenomenal reports by observers when looking at these monocular stimuli. The subjective contours in Figure 2 almost always are seen as bounding an opaque figure which is *in front of* the other pattern elements. This perception of relative depth is quite strong, as may be anecdotally illustrated by the fact that one observer, while viewing a pattern similar to Figure 2A, accused the experimenter of cutting a triangle out of a piece of white paper, different than the background, and pasting it onto the stimulus card. If the important aspect of the stimulus display is the displacement of one part of the array in depth, then one ought to be able to demonstrate the ability of other depth cues to produce the perception of subjective contours.

Shading and Subjective Contours

One such depth cue which has long been known to produce subjective planes varying in apparent distance from the background is the cue of shading (Brunswik, 1935). Figure 3 shows a stimulus configuration which utilizes this cue. In this figure the word FEET is clearly seen in white letters which stand out from the background. It is difficult to believe, when looking at this figure, that about 50% of each letter is not present in the stimulation. Most observers report the perception of white letters, whiter than the background, with the contours defining the letters faintly visible. These contours are subjective contours in the same way that the contours in Figure 2 were. The boundaries defining the edge between the white of the raised plane of the letters and the white of the background are subjectively created from homogeneous stretches of the stimulation. That this phenomenon is due to the

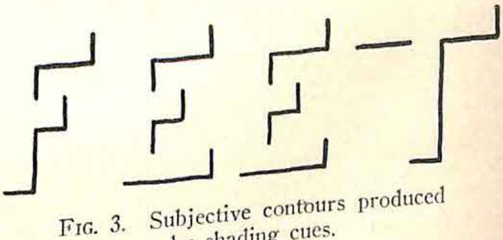


FIG. 3. Subjective contours produced by shading cues.

presence of the depth cues and not simply to the presence of these particular lines in the field is shown in Figure 4.

Figure 4A shows a letter H created in the same manner as the letters in Figure 3. Here the shading provides a monocular depth cue sufficient to create an apparent plane displaced forward in space, and the contours which bound it are visible. Figure 4B contains the same lines as Figure 4A, except that they have been rearranged so that the shading cues are inconsistent. Thus, a variety of "impossible figure" with locally consistent and globally inconsistent depth has been created, such as the figures explored by Hochberg (1968). Most observers report that Figure 4B is much harder to see as separate from the background, and the subjective contours are never very clear or stable. They tend to spontaneously appear and disappear. As one observer put it, the subjective contours of Figure 4B lack the "reality" of the contours of Figure 4A.

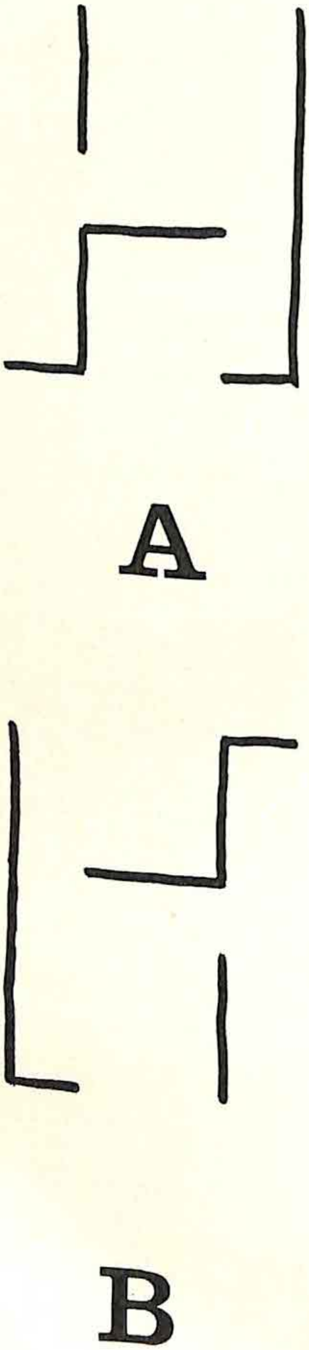


FIG. 4. A subjectively contoured letter H produced by shading cues (A). The same shading lines, arranged to produce inconsistent cues, produce an unstable subjective contour (B).

The perceptual creation of three-dimensional planes out of a two-dimensional array of elements is not a new phenomenon. Hochberg and Brooks (1960) have shown that when a complex two-dimensional figure is presented to observers, they very frequently "simplify it" by coding it as a three-dimensional figure. Thus the meaningless, asymetrical jumble of two-dimensional lines in Figure 4A is more compactly coded as a single three-dimensional letter H. The main difference between these figures and the Kopferman cube figures used by Hochberg and Brooks is that in these configurations the subject not only renders the percept into one of tridimensionality, but supplies the missing edges to make the stimulus apparently complete.

Interposition as a Cue for Subjective Contour

Now return to the stimuli presented in Figure 2. Since it has already been shown that a monocular depth cue, namely shading, can act in the same fashion that the binocular disparity cue does to form planes and contours out of homogeneous stimulus regions, a new way of looking at these subjective contour configurations has emerged. Given the fact that the triangle in Figure 2A is clearly seen in front of the other elements, let us look for the relevant depth cues which cause this displacement and the corresponding creation of the subjective edges. The depth cue becomes quite apparent when the observer's descriptions of the figure are recalled. It was described as a white triangle *in front of* an inverted black outline triangle, with its corners *resting on* three black squares. The depth cue involved in this pattern is therefore probably interposition. If there were an opaque white triangle overlying a black outline triangle and three black squares, the same retinal pattern would be produced, except, of course, for the existence of contours and apparent brightness or color differences in the field. Thus the observer creates the surface indicated by the depth cues and edges it with subjectively created contours to bound it. Interposition cues can also be found in the remaining patterns in Figure 2. Figure 2B is easily seen as a white rectangle interposed in front of the intersection of two black lines, its corners resting on four black circles, and Figure 2C is seen as a white circle resting on the intersection of two heavier black bars. Figure 2D provides the same interposition cues as are found in Figure 2A, except that the brightnesses are reversed, such that the interposition cues indicate that the figure which must be present is a black triangle.

A very interesting demonstration of the presence of depth cues in these configurations is shown in Figure 5. This figure is a variant of Figure 2A where a circle on the apparent triangle and one beside it on the background have been placed. Now if a cue for depth is present, then a situation arises where two stimuli both subtend the same visual angle, but one is seen as apparently closer than the other. If the depth cue is strong enough to evoke the operation of the constancy scaling mechanism, then the circle on the apparently closer plane (the triangle) should be seen as smaller than the circle on the apparently more distant plane (the background). A copy of this figure was presented to 12 subjects. Each subject was required to judge which of the two circles appeared larger, using a forced-choice technique. Eleven of the 12 reported that the circle on the subjective triangle appeared smaller than the one outside it ($p < .01$, sign test). These results seem to indicate not only the presence of the interposition cue in these configurations, but the fact that they are relatively strong. It should be clear that this demonstration is lacking in a control group, since the effect might also be due to configurational properties. Simply adding a real contour does not suffice as a control, since then both the interposition cues for depth and configurational properties would still be present. Removing the interposition cues on the other hand, drastically changes the configuration. Suffice it to say that accompanied with the reported depth difference between the subjectively seen contour and the background, these results are clearly in accord with the hypothesis that depth cues are present in configurations which produce subjective contours.

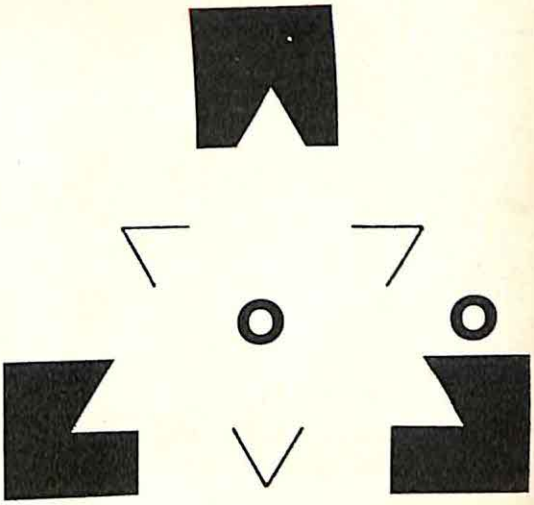


FIG. 5. Both circles subtend the same visual angle; however, the apparent depth difference produces an apparent size difference.

Figure 6 is an example of how the explicit use of the interposition cue can result in the perception of subjective contours. This figure has been adopted from one presented by Matthaei (1929). When shown this figure, the immediate response of most observers is that they are seeing a word which is *partially blocked by a white strip which has been overlayed*. In fact the principle guiding the construction of this figure was to simply remove those portions of the display which would not be seen if such an opaque strip were physically present.

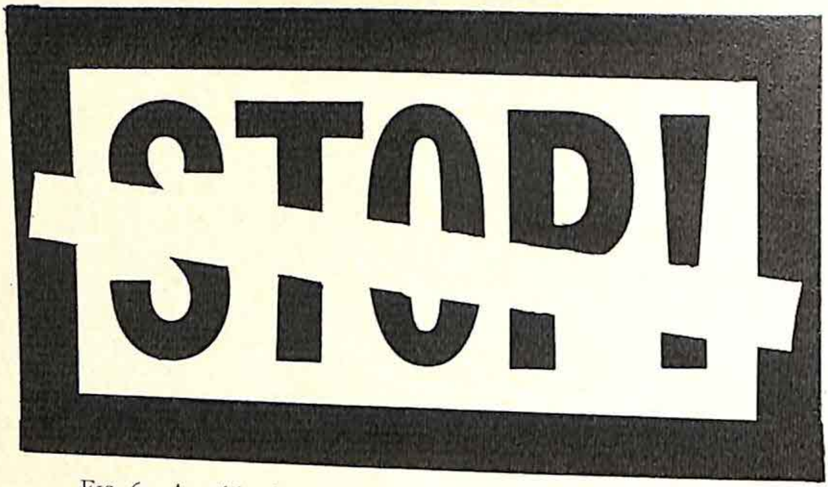


FIG. 6. A subjective contour produced by interposition cues.

Texture and Perspective

If the author's analysis is correct, and it is the presence of depth cues which are constant with the presence of forms or planes at various depths, then almost any monocular depth cue could be utilized to produce the perception of subjective contours. The only prerequisite is that cues be strong enough so that the configuration is seen as tridimensional rather than as bidimensional. This has been done for a number of these cues, and Figure 7 presents a combination of texture and perspective cues which demonstrates the effect very clearly. This configuration is patterned after the stimuli presented by Gibson (1950). The general impression produced by this figure is that of two planes differing in relative depth. The plane in the upper part of the figure is perceived as more distant than the plane in the lower half of the figure. There is apparently a sharp drop or a cliff separating these two planes. The contour delineating the edge of the cliff is clearly seen. It is visible as a sort of whitish line. This light line is, of course, not present in the physical stimulation. As in the previous stimuli, it bounds the edge between two planes seen at different depths.

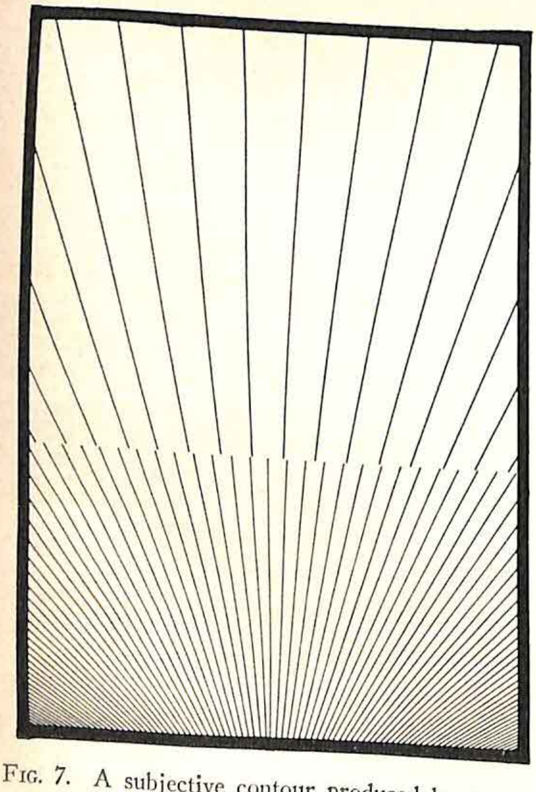


FIG. 7. A subjective contour produced by texture and perspective cues (after Gibson 1950).

Closure as an Explanatory Mechanism

At the outset it was noted that the most frequently offered explanation for subjective contours is that they are an example of the Gestalt principle of closure (Kanizsa, 1935; Lawson & Gulick, 1967; Pastore, 1971; Rubin, 1921; Schumann, 1904). It is interesting to compare these effects which are being ascribed to the effects of depth cues with the perceptual effect generally called closure. Figure 8A is a configuration which utilizes interposition cues to produce the perception of a white triangle with its corners resting on three black circles. The phenomenological impression is that this triangle is in front of the background and the other pattern elements, and that it is bounded by contours which are faintly visible to the observer. Figure 8B is a typical figure which is associated with the phenomena of closure. This figure is easily perceptually completed and recognized by the observer as a triangle. The percept which results from this configuration is quite different from that of 8A. The figure does not lie in front of the plane and the figure is still seen as being incomplete. Most importantly, this figure is not bounded by subjective contours which clearly set it off from the background. It is thus clear that by degrading the stimulus in such a way that the relative depth cues are not readily available then closure alone is not sufficient to produce the perception of subjective contours. Figure 8B is also interesting in that it clarifies a major criterion for the perception of the subjective contour. Discussions with Julian Hochberg have suggested that a depth cue be defined as some aspect of a configuration which can be read as consistent with a given spatial arrangement of objects at different relative distances. To the extent, then, that there is no evoked perception of objects differing in relative depth in Figure 8B, there are no depth cues in this pattern. However, since the definition of a depth cue includes any configurational change which may result from objects in depth relative to each other, then there are potential depth cues in Figure 8B. Any time there is a discontinuous black line on a white ground, the discontinuity may be caused by an interposed white figure. Thus if one looks at Figure 8B and imagines that it is a complete black outline triangle, then one can begin to configure a white object interposed in front of it. This figure may either be an amorphous white cloud or, more spectacularly, an inverted white subjective triangle. Such an inverted triangle is not generally seen, although when alerted to its presence an observer may be able to make out its form. This illustrates that the mere presence of potential depth cues in an array is not sufficient to evoke the perception of a subjective contour. The reason for its appearance in some situations and not in others is sug-

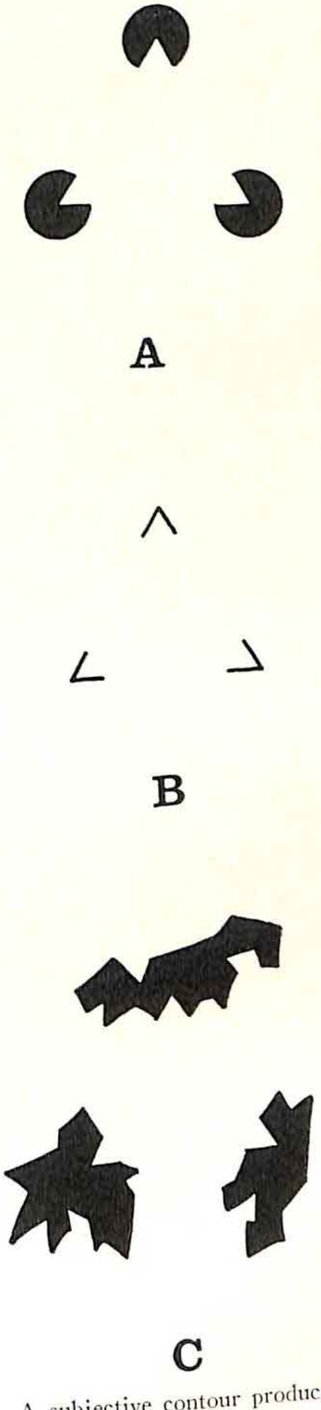


FIG. 8. A subjective contour produced by interposition (A). A typical closure configuration which produces no subjective contour (B). Although corner elements are still present, this configuration generally yields no subjective contours (C).

gested from the work of Kopfermann (1930) and Hochberg and Brooks (1960), who showed that complex two-dimensional line figures were seen as tridimensional if such a percept resulted in a simpler overall organization. Thus Figure 8A is simplified from a meaningless pattern of irregular forms into an arrangement in depth of four simple symmetrical forms. Figure 8B can be simplified into a triangular organization without the necessity of rendering the percept into depth. Hochberg (1970) has shown several convincing demonstrations of the effect of expectation on perceptual organization based on the utilization of local and global depth cues. Thus in Figure 8B, once the observer is alerted to the presence of the depth cues and the possibility of an object in depth, the perceptual organization may change in such a manner as to allow for the perception of the subjective contours. Figure 8C is another form of the same figure and makes a similar point. Here the corner elements are still present and we also have the black masses associated with Figure 8A. If subjects are shown this figure, they usually describe three black nonsense figures (more or less meaningful according to their personal styles). The spontaneous report of a white triangle which is invariably obtained in Figure 8A is missing. By reducing the regularity of the forms at the corners of the subjective figure, the likelihood that the configuration will be perceptually simplified by assuming a white triangle interposed in front of the background has been reduced. It is interesting to note that when the white triangle is pointed out in Figure 8C, many observers report that it is faintly visible, as are the subjective contours bounding it.

Conclusions

These demonstrations may be summarized then by indicating that all of the figures which have been presented that contain subjective contours are seen as associated with apparent surfaces or planes that are in depth relative to the remainder of the array. These planes are perceptually created on the depth cues present in the stimulus. Since every plane must have an edge, the bounding contour is supplied by the perceptual system. This contour thus segregates the homogeneous stimulation into discrete units, one comprising the plane in depth and the other, the background stimulation. The edges of the plane are subjective contours, since there is no discontinuity in physical stimulation at the locus where they are seen in space. In this way a chaotic collection of complex two-dimensional elements has been simplified into a simple and easily coded three-dimensional array of meaningful or symmetrical elements. Thus, a subjective contour is simply the edge of a subjective plane, and a subjective plane is a surface which ought to be present on the basis of the available depth cues, but is not except in the mind of the perceiver.

REFERENCES

BRUNSWIK, E. *Experimentelle psychologie in demonstrationen.* Vienna: Springer, 1935.

GIBSON, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.

HOCHBERG, J. In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception.* New York: Holt, Rinehart & Winston, 1968.

HOCHBERG, J. Attention, organization, and consciousness. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis.* New York: Appleton-Century-Crofts, 1970.

HOCHBERG, J., & BROOKS, V. The psychophysics of form: Reversible-perspective drawings of spatial objects. *American Journal of Psychology,* 1960, **73**, 337–354.

JULESZ, B. Binocular depth perception without familiarity cues. *Science,* 1964, **145**, 356–362.

KANIZSA, G. Marzini quasi-percettivi in campi con stimolazione omogenea. *Rivista di psicologia,* 1955, **49**, 7–30.

KAUFMAN, L. On the nature of binocular disparity. *American Journal of Psychology,* 1964, **77**, 393–402.

KAUFMAN, L., & PITBLADO, C. Further observations on the nature of effective binocular disparities. *American Journal of Psychology,* 1965, **78**, 379–391.

KOPFERMANN, H. Psychologische Untersuchungen über die wirkung Zweidimensionälar Darstellungen korperlicher Gebilde. *Psychologische Forschung,* 1930, **13**, 293–364.

LAWSON, R. B., & GULICK, W. L. Stereopsis and anomalous contour. *Vision Research,* 1967, **7**, 271–297.

MATTHAEI, R. *Das Gestalt problem.* Munich: Bergmann, 1929.

OSGOOD, C. E. *Method and theory in experimental psychology.* New York: Oxford University Press, 1951.

PASTORE, N. *Selective history of theories of visual perception 1650–1950.* New York: Oxford University Press, 1971.

RUBIN, E. *Visuell wahrgenommene figuren.* Copenhangen: Gyldendal, 1921.

SCHUMANN, F. Einige Beobachtungen über die Zusammenfassung von Gesichtseindrücken zu Einheiten. *Psychologische Studien,* 1904, **1**, 1–32.

SHIPLEY, T. Visual contours in homogeneous space. *Science,* 1965, **150**, 348–350.

WHITE, B. W. Stimulus-conditions affecting a recently discovered stereoscopic effect. *American Journal of Psychology,* 1962, **75**, 411–420.

(Received December 2, 1971)

Psychological Review
1972, Vol. 79, No. 4, 368

LETTERS TO THE EDITOR

A POSSIBLE TEST OF HEBB'S HYPOTHESIS CONCERNING IMAGERY

Hebb (1968) has suggested that eidetic imagery is a function of first-order cell assemblies, which he equates with the "simple" cells reported by Hubel and Wiesel (1968). He also suggests that the more usual type of imagery is a function of higher order assemblies.

It has been claimed (Day, 1962) that the negative aftereffect to a tilted line is mediated by excitatory and inhibitory processes in the visual cortex at the level of Hubel and Wiesel's simple cells. If this hypothesis is accepted, then there is obtainable evidence which will demonstrate the truth or falsity of Hebb's suggestion.

If an eidetic imager images a tilted line, he should, according to Hebb, reinstate the activity of the first-order cell assemblies (i.e., those involved in the perception of such contours). He should then subsequently misjudge the position of a test line, just as if he had initially perceived the line. If this occurred, it would confirm Hebb's hypothesis. If, given suitable controls, it does not occur, then it would suggest that the locus of eidetic imagery is not that of first-order cell assemblies.

If doubts existed about the suitability of the tilt aftereffect as a variable, any aftereffect mediated by known analyzers at the level of simple cells would be acceptable.

REFERENCES

Day, R. H. Excitatory and inhibitory processes at the basis of contour shift and negative aftereffect. *Psychologia*, 1962, **5**, 185–193.

Hebb, D. O. Concerning imagery. *Psychological Review*, 1968, **75**, 466–477.

Hubel, D. H., & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 1968, **195**, 215–243.

Roderick P. Power
Department of Psychology
Queen's University of Belfast
Northern Ireland

(Received February 9, 1972)

This is an interesting possibility. Day's effect was obtained with relatively simple stimulus conditions, whereas the eidetic effect has always been obtained with a meaningful picture, but it should be possible to design one in which a clearly marked sloping line is a component, and then make the test. The meaningfulness may be significant, since higher order assemblies may be involved in maintaining the first-order activity.

Donald O. Hebb
Department of Psychology
McGill University
Montreal 101, Quebec

DEVELOPMENTAL PSYCHOLOGY

Editor: Boyd R. McCandless

Developmental Psychology contains articles which represent the broad range of growth and development and their associated variables. Chronological age as well as sex, socioeconomic status, and effects of physical growth variables (including time of maturation and body build) are all considered developmental variables.

Although the adolescent and the aging population appear to deserve particular attention at this time, *Developmental Psychology* indeed covers the entire span. Cross-species articles and articles concerning developmental research with retardates are also included.

The articles in *Developmental Psychology* are not limited to any one approach such as "experiments versus studies" or "cognition versus motivation."

Bimonthly: Two volumes per year
Subscription Price: $20.00 per year

AMERICAN PSYCHOLOGICAL ASSOCIATION
1200 Seventeenth Street, N.W.
Washington, D.C.
20036

2 FEB 1974

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

announces new developments in

Psychological Abstracts Information Services

The APA has been expanding the *Psychological Abstracts (PA)* information service offered to the psychological community since *PA*'s inception in 1927 and NOW offers the capability for automated search, in a matter of minutes, of all records published in *PA* over the last 5 years. This new development may revolutionize the researcher's use of the psychological literature. Search and retrieval is now as easy as 1, 2, 3 . . .

1 BEGIN SEARCH.

2 REQUEST: *PSYCHOLOGICAL ABSTRACTS* INFORMATION SERVICES◣

3 8 ENTRIES. DO YOU WANT TO PROCESS ANSWERS: NO, PRINT, OR MODIFY?

4 MODIFY – AUTOMATED SEARCH AND RETRIEVAL SERVICES◣

5 3 ENTRIES. DO YOU WANT TO PROCESS ANSWERS: NO, PRINT, OR MODIFY?

6 PRINT◣

7 ENTER DESIRED OUTPUT.

8 FULL SUMMARY◣

9 PASAR (*PA* SEARCH AND RETRIEVAL), A CENTRALIZED AUTOMATED SERVICE THAT PROVIDES RESPONSE TO A FORMULATED PROBLEM OF ANY DEGREE OF SPECIFICITY.

10 PASAR CHARGE = $15 PROCESSING FEE + $2.25 PER MIN. COMPUTER TIME (AVE. = $40-$60; MAXM. = $95).

11 PADAT (*PA* DIRECT ACCESS TERMINAL), A SERVICE THAT ALLOWS REQUESTOR TO INTERACT WITH THE DATA BASE THROUGH A REMOTE COMPUTER TERMINAL IN HIS OWN FACILITY.

12 PADAT CHARGE = $25 FEE SUBSCRIPTION PER MO. + $1 PER MIN USE TIME (MINM. 1 HR. USE TIME PER MO.).

13 PATELL (*PA* TAPE EDITION LEASE OR LICENSING), A SERVICE THAT MAKES *PA* MACHINE READABLE RECORDS AVAILABLE TO INSTITUTIONS TO OPERATE THEIR OWN SEARCH AND RETRIEVAL SERVICES.

14 PATELL CHARGE = $3,000 FOR CURRENT YR., $2,000 FOR PREVIOUS YR., $800 FOR OTHER PREVIOUS YR.

For full information retrieval, write to:

PAIS
AMERICAN PSYCHOLOGICAL ASSOCIATION
1200 Seventeenth Street, N.W.
Washington, D.C. 20036

Those who would like to see PADAT in operation may request that a demonstration be arranged for their group at APA Central Office headquarters, or other negotiable locations.

Ack
Jacket 2 issue
to the Library [illegible]
R D as
1/1/74

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 79, No. 5 September 1972

Articles

Complexity and the Representation of Patterned Sequences of Symbols
Herbert A. Simon 369

Dual Functions of Consciousness
Tim Shallice 383

Reinforcement, Expectancy, and Learning
Robert C. Bolles 394

Processing Dimensional Stimuli: A Note
G. R. Lockhead 410

Theoretical Notes

On Reasoning, Congruence, and Other Matters
Janellen Huttenlocher and E. Tory Higgins 420

On the Evidence Concerning J. Huttenlocher and E. T. Higgins' Theory of Reasoning: A Second Reply
Herbert H. Clark 428

Editorial Note 432

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

HENRY ALKER
D. E. BERLYNE
DONALD T. CAMPBELL
R. G. CROWDER
MORTON DEUTSCH
HOWARD E. EGETH
STEPHEN S. FOX
W. R. GARNER
RICHARD L. GREGORY
JEANNE GULLAHORN
RONALD L. KOTESKEY
RICHARD S. LAZARUS
SAMUEL LEINHARDT
MARVIN LEVINE
JOHN T. MACNAMARA
GEORGE A. MILLER
THEODORE M. NEWCOMB
MICHAEL POSNER
ROBERT A. RESCORLA
SANDRA SCARR-SALAPATEK
ROGER SHEPARD
RICHARD F. THOMPSON
SUSAN ERVIN TRIPP
ELLIOT S. VALENSTEIN
HANS WALLACH
JOACHIM F. WOHLWILL

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts submitted to the *Psychological Review*. In each issue the Editor will list in the *Review* those editorial consultants who have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowledged.)

The *Psychological Review* is devoted to articles of theoretical significance to any area of scientific endeavor in psychology. Except for occasional articles, manuscripts exceeding 7,500 words will be discouraged. Ordinarily, manuscripts which consist primarily of original reports of research or of reviews of the literature should be submitted to other journals. Methodological and statistical articles are not appropriate for the *Psychological Review*. See the editorials in the *Review* (1965, **72**, 1–2, 1971, **78**, 1–2) for further details.

All manuscripts must be submitted in triplicate. In reference lists, give journal titles in full; do not abbreviate. All copy must be double spaced, including the references. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints; duplicate figures may be photocopies or pencil-drawn copies.

Beginning with the 1971 issue, Letters to the Editor will be published on an experimental basis. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obtained from the APA Central Office.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Manuscripts should be addressed to the Editor, George Mandler, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of materials copyrighted by the American Psychological Association: Written permission must be obtained for copying or reprinting any tables, figures, or text of more than 500 words in length. Persons or organizations who obtain permission to reproduce any such copyrighted materials must include on the first page of the reproduced materials the exact copyright notice which appeared on the Association's original publication. There will be a charge of $10.00 per page for such permission except when the requester is the author of the article to be reprinted, or when the article is to be reproduced in a limited number of copies solely for instructional purposes.

HAROLD P. VAN COTT
Managing Editor

LORRAINE BIROS
Technical
Editor

HENRY S. SALMON
Advertising and Business
Manager

ANNE REDMAN
Subscription Manager

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1200 Seventeenth St., N. W., Washington, D. C. 20036

$10.00 volume

$2.00 issue

Second-class postage paid at Lancaster, Pa.

Send all communications, including address changes, to 1200 Seventeenth St., N. W., Washington, D. C. 20036. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PSYCHOLOGICAL
REVIEW

Diary No 727(6)
Date 92
d. N Ci6
Bureau E. I. Psy. Research.

Vol. 79, No. 5 September 1972

COMPLEXITY AND THE REPRESENTATION OF PATTERNED SEQUENCES OF SYMBOLS[1]

HERBERT A. SIMON[2]

Carnegie-Mellon University

All of the principal alternative theories that explain behavior in tasks involving patterned sequences are variants on a basic theme. They agree in proposing that subjects perform these tasks by inducing pattern descriptions from the sequences, and that these descriptions incorporate the relations of "same" and "next" between symbols, iteration of subpatterns, and hierarchic phrase structure. As a result, measures of the relative complexity of different patterns show high agreement among the theories, and it becomes difficult to choose among alternative encoding schemes as explanations of the human performance.

In the pages of the *Psychological Review* alone, at least four apparently distinct theories have been proposed during the past decade to explain how human subjects process patterned sequences of symbols (Leeuwenberg, 1969; Restle, 1970; Simon & Kotovsky, 1963; Vitz & Todd, 1969). Additional theoretical proposals have appeared in other journals and books (e.g., Feldman, 1959, 1961; Feldman, Tonge, & Kanter, 1963; Glanzer & Clark, 1962; Gregg, 1967; Klahr & Wallace, 1970; Laughery & Gregg, 1962; Pivar & Finkelstein, 1964; Simon & Sumner, 1968; Williams, 1969). These theories differ radically from earlier explanations (in the traditions of S–R theory or stochastic learning theory) by postulating specific encodings of the sequences and by predicting from the hypothesized encodings a number of characteristics of subjects' behavior (e.g., loci of their errors, the relative difficulty of different problems, judged complexity of patterns, and so on).

In this paper it will be shown that these various theoretical proposals, so different in their surface appearance, derive from a common central core, and that it is this commonality that accounts for their success in explaining behavior. Our analysis requires us to pay careful attention to the notations in which the patterns of sequences are described (by subjects and by theorists); hence, the analysis may cast some light on the general psychological significance and implications of notational devices.

Representational (or "information-processing") explanations of serial pattern processing generally postulate some kind of language or notation[3] for expressing or de-

---

[1] This work was supported in whole or in part by Public Health Service Research Grant MH-07722 from the National Institute of Mental Health. I am grateful to William G. Chase, Lee Gregg, David Klahr, Kenneth Kotovsky, Allen Newell, and Paul C. Vitz for comments on an earlier draft.

[2] Requests for reprints should be sent to Herbert A. Simon, Department of Psychology, Carnegie-Mellon University, Schenley Park, Pittsburgh, Pennsylvania 15213.

[3] I shall not distinguish systematically between the terms "language" and "notation," but will generally use the former to refer to the broader features of a representation, and the latter to refer to the more particular features of a representation.

© 1972 by the American Psychological Association, Inc.

scribing the patterns, and one or more processes for inducing pattern descriptions from sequences, for storing descriptions, for reproducing sequences, and for extrapolating them. Not all of the theories incorporate all of these elements.

The theories have been applied to various kinds of sequential material, including binary sequences (Feldman, 1959, 1961; Glanzer & Clark, 1962; Restle, 1967), sequences of letters and/or digits (Restle, 1970; Simon & Kotovsky, 1963; Vitz & Todd, 1969), linear geometric figures (Leeuwenberg, 1969), intelligence test items (Williams, 1969), and musical scores (Restle, 1970; Simon & Sumner, 1968).

The tasks that have been used to test the various representational theories differ not only in the kinds of patterned sequences they use, but also in what they ask the subject to do. In some cases, he must remember and reproduce a finite (or cyclically recurring) sequence (e.g., Gregg, 1967; Laughery & Gregg, 1962; Restle, 1967, 1970; Vitz & Todd, 1969); in others, he must extrapolate a partial sequence—that is, predict its continuation (e.g., Feldman, 1959; Simon & Kotovsky, 1963; Williams, 1969); in still others, he must rank different sequences by complexity (e.g., Leeuwenberg, 1969; Vitz & Todd, 1969).

Several of the basic elements that are incorporated in the representational languages can be found in work that predates these theories. Runs of identical symbols, for example, and simple and double alternations were discussed by earlier theorists (see e.g., Estes & Burke, 1955; Goodnow & Pettigrew, 1955). I limit this comparison, however, to contemporary information-processing analyses of the phenomena.

INFORMATION PROCESSING AND INFORMATION THEORY

Information theory provides one direct method for measuring the complexity of sequences, by equating complexity with amount of information. Among the approaches under discussion here, the theory of Vitz and Todd (1969) is based on this information-theoretic measure of complexity.

The amount of information in a patterned sequence, in information-theoretic terms, is equal to the number of symbols required to describe the sequence when the maximally efficient code is used. If the coding alphabet consists of the binary digits, 0 and 1, then the information content of a sequence is equal to the number of such digits, or bits, in its code when it is coded efficiently. What code will be most efficient for representing a set of sequences depends both on the collection of sequences that belong to the set and on the relative frequency with which they occur. An encoding can be made more efficient for a given set of sequences by assigning short codes to sequences that occur frequently.

Suppose, for example, that five sequences have been encoded by the 12-symbol binary strings: 111100001111, 111111100000, 110000110000, 111111111111, and 010101010101, respectively. Suppose that the first and fourth of these sequences occur somewhat more frequently than the second and third, and that the fifth is exceedingly rare. Then we can encode them more efficiently by making the transformations: 1111 → 111, 0000 → 110, 111 → 101, 000 → 100, 11 → 011, 00 → 010, 1 → 001, and 0 → 000. With this scheme the first four sequences are encoded: 111110111, 111101110000, 011110011110, and 111111111, respectively. The second and third sequences have codes of length 12, as before, but the first and fourth now have codes of only length 9. If complexity is measured by code length, then in the original encoding, all the sequences were of the same complexity; while in the new encoding, the first and fourth are simpler than the second and third. But while these particular four sequences all have codes of the same or shorter length in the new encoding as compared with the original one, this is dramatically untrue of the fifth sequence. The pattern 010101010101, also 12 symbols in length, would be recorded as 000001000001000001000001000001000001—a string of 36 symbols. The price paid for shortening the codes of some sequences is to lengthen the codes of others. The total number of encodings that can be represented by binary strings of length n is exactly 2^n, independent of the code used. The recoding will nevertheless increase average code efficiency if the fifth sequence is sufficiently rare.

Thus, information theory does not provide us with an unambiguous index of sequence complexity, but only measures complexity relative to some particular code. The complexity rankings of the sequences in a set can be changed at will by altering the encoding scheme on which the index of complexity is based. If an index of complexity is to have significance for psychology, then the encoding scheme itself must have some kind of psychological basis.

Encoding Alphabet

Also, it is not obvious why, in measuring code lengths, the binary alphabet, consisting only of the symbols 0 and 1, should be employed, rather than richer alphabets —the 10-decimal digits, for example, or the digits supplemented by the letters of the Roman alphabet. If we measure complexity by code length in *symbols* instead of code length in *bits*, then the complexity of sequences can be reduced by using a larger coding alphabet, each symbol of which can convey several bits of information. The alphabet of digits allows more than three bits to be encoded per symbol ($10 > 2^3$), while the Roman alphabet can encode between four and five bits per symbol ($2^4 < 26 < 2^5$).

Psychological theory provides good reasons why code length should be measured in terms of symbols rather than bits. If we replace the term "symbol" by "chunk," then we see that this observation is at the root of Miller's (1956) chunking hypothesis. The chunking hypothesis asserts that human short-term memory is limited in the number of symbols (or chunks) it can hold, not in the number of bits of information it can hold. Similarly, the EPAM (elementary perceiver and memorizer) theory of verbal learning asserts (Simon & Feigenbaum, 1964) that fixation of information in human long-term memory requires a constant time per symbol (or chunk), and not a constant time per bit.

Of course, when measuring complexity by code length, we cannot admit arbitrarily large coding alphabets, for these would allow us to reduce complexity to any desired level. In the limit, we could assign one letter of the alphabet to each sequence in the population to be encoded, with the result that each sequence would have a complexity of exactly one symbol. The solution to this dilemma, derivable from the EPAM theory, is to admit anything as encodable in a single symbol that has become familiar through prior learning. Anything recognizable by a subject as a "chunk," as the result of previous training or experience, is assumed to be codable into a symbol. The available coding alphabet then consists of the set of such symbols.

That the complexity of sequences must be measured relative to the psychological characteristics of the subject can be demonstrated in a variety of ways. Perhaps the most trivial is to observe that literate persons in a culture that uses the Roman alphabet will judge the sequence A B C D E to be simpler than the sequence A C E D B. Illiterate persons, and monolingual Chinese and Russians may judge the two sequences to be equally complex. Clearly, then, the presence or absence of a particular alphabet stored in long-term memory can have a major influence on judgments of complexity.

Code Length and Pattern Complexity

The theory to be proposed here takes code length as the basic measure of sequence complexity. It is postulated that most subjects in our culture will use approximately the same coding alphabets and encoding procedures. The code length of a sequence, then, means the number of symbols in the encoding when the alphabets and coding procedures common to the culture are used. We may call this encoding the *common pattern* of the sequence and we will measure the complexity of the sequence by the length, in symbols, of this pattern.

From this point of view, any particular pattern language implies a psychological (or sociological) assertion about the alphabets and coding procedures that a subject will use when instructed to process certain sequences. While subjects' pattern language is not directly observable, it has observable consequences. If two or more pattern languages predict different complexity orderings of the sequences in some set, then empirical observations of the relative difficulty that the subject experiences with members of the set allow, at least in principle, a choice to be made of the language that corresponds most closely to the actual encoding scheme he is using. This is the underlying logic of virtually all experiments using patterned sequences as stimuli.

Now, in fact, there is almost always a high correlation among different measures of pattern complexity when several such measures are applied to the same set of patterns. Since the presence of correlation is not logically necessary, it must result from the fact that the coding schemes actually invented by investigators share some basic common features, which also lie at the base of the coding schemes used by subjects.

The next section discusses the common elements that are incorporated in the coding schemes of Simon and Kotovsky (1963), Glanzer and Clark (1962), Simon and Sumner (1968), Leeuwenberg (1969), Vitz and Todd (1969), and Restle (1970), as well as the specific differences among these schemes. These six papers represent fairly well the range of alternative coding schemes that have appeared in the literature.

PATTERN LANGUAGES

In serial tasks, both terminating and nonterminating sequences have been used. To simplify the treatment, regard all sequences as nonterminating. Thus, a sequence of finite length (e.g., 12321) will be assumed to repeat indefinitely (123211232112···). As is customary in programming notation, the asterisk, *, will be used as superscript to denote indefinite repetition: $(12321)^*$.

Conventional and Descriptive Pattern Names

A sequence of symbols can be designated by a conventional name or by a descriptive name. (We call the latter the *pattern* of the sequence.) Thus "Fibonacci Numbers" is the conventional name of the sequence: 1 2 3 5 8 13 21 34 ···. This sequence can be named descriptively as "the sequence each of whose members is the sum of the two preceding members." A descriptive name for a sequence, its pattern, provides the information required by an appropriate encoder to extrapolate the sequence indefinitely. A conventional name does not provide this information; hence if only the conventional name is communicated to a decoder, the latter must already have stored the description (or its informational equivalent) if it is to be able to extrapolate the sequence.

If a pattern consisting of a finite string of symbols is to describe an infinite sequence, then the pattern must define one or more relations between each element of the sequence and preceding elements, or between each element and its order number in the sequence. An example of a pattern of the first kind is provided by the familiar formula for the Fibonacci Numbers: $F(i) = F(i-1) + F(i-2)$. An example of a pattern of the second kind is provided by a formula for the sequence of squares, 1, 4, 9, 16 ···: $S(i) = i^2$.

A pattern, then, is a finite string of symbols that states the rule governing the indefinite continuation of a nonterminating sequence. The symbols of the sequence are numbered in order, beginning with 0. The pattern must express relations between the ith symbol as the dependent variable, and one or more preceding symbols and/or constant symbols as independent variables. Thus, the language must be capable of expressing the relations themselves, the names of the symbols that enter as dependent variables, and the names of the symbols that enter as independent variables.

Consider, by way of example, the simple sequence:

$$\text{ABCD} \cdots. \qquad [1]$$

In the pattern to describe a sequence, the letters of the alphabet employed in the sequence can represent themselves. The letter n will represent the relation of *next* on an alphabet. The ith symbol in the sequence will be represented by Si; and the relations of sum and difference of pairs of integers, by $+$ and $-$, respectively. Using these notational devices, a pattern for Sequence 1 is

$$S0 = \mathrm{A};\ Si = n(S(i-1));\ i(1:{}^{*}) \qquad [2]$$

where the last relation indicates that i is to range from 1 to infinity. Counting all punctuation marks as symbols, this pattern has length 24, which number we may therefore take as a measure of the complexity of Sequence 1 in this encoding.

Let us consider, in turn, each of the components of such an encoding scheme: the relations it admits, the way in which it names symbols, and the way in which it notates iteration.

Relations

What relations are available for defining the pattern of a sequence depends on the properties of the set of symbols, S, from which the sequence is constructed. For any set of symbols, S, we can define the identity relation (s, for "same") on the product set $S \times S$. Using this relation, we can describe simple patterns. Thus the pattern:

$$\mathrm{AQAQAQ} \cdots. \qquad [3]$$

can be defined by:

$$S0 = \mathrm{A};\ S1 = \mathrm{Q};\ Si = s(S(i-2));\ i(2:{}^{*}) \qquad [4]$$

(This pattern can also be defined simply by: $S(2i) = \mathrm{A};\ S(2i+1) = \mathrm{Q};\ i(0:{}^{*})$.)

If the symbol set, S, is ordered, that is, constitutes an alphabet, then we can also define on $S \times S$ the successor relation (n, for "next"), and its inverse (p, for "immediate predecessor"). Moreover, these two relations, together with s, are the only relations that are common to all ordered sets.[4] With two exceptions, discussed below, they are also the only relations that enter into the serial pattern tasks that have been used in psychological experiments and tests.

The relations s and n are central to the pattern language developed by Simon and Kotovsky (1963) and extended by Simon and Sumner (1968) and by Williams (1969). Both play an equally central role in Leeuwenberg's (1969) pattern language[5] and in Restle's (1970). Restle calls the s relation r (for "repeat") and the n relation t (for "transpose").

Glanzer and Clark (1962) and Vitz and Todd (1969) limit themselves to binary and ternary sets of symbols (the sets 0,1; a,b; and a,b,c), hence make no use of the n relation, but only of the s relation. Therefore, their pattern languages are both simpler and more limited in range of application than the languages used by the other authors.

We come now to two additional kinds of relations that are used by Simon and Sumner (1968; and subsequently by Restle, 1970) to describe certain patterns—relations that are relevant only to the alphabet of integers. With the integers, we may use first of all, the relations of *sum* (+) anc *difference* (−). These can be defined ir terms of the successor and predecesso relations, as in the Peano axiomatization o arithmetic (e.g., $(i+2) = n(i+1) = n^2(i)$ $(i-1) = p(i)$), but it is often more convenient, notationally, to retain the arithmetic relations as well as the successor anc predecessor relations.

Finally, with the alphabet of integers, w may also define the *complement* relatio (denoted ck, where k is an integer). To se what this means, consider the sequence:

$$123321 \qquad [5$$

---

[4] More precisely, all other relations that are common to all ordered sets can be defined in terms of these, for these are the only relations required for the definition of an ordering.

[5] The resemblance of Leeuwenberg's langua to the others is not easy, at first, to detect unde neath his notation. An integral sign plays t role of n in his language, and parentheses t role of s. Thus, the pattern $\int(1)$ describes t sequence 1234 ··· in his notation. That is say, (1) denotes $(1)^*$, and the integral sign desi nates successive summations: 1, 1+1, 2+1, e In the notation of Simon and Kotovsky (196 this same sequence would be encoded as [$\mu \leftarrow$ this same sequence would be encoded as tegers] $(n(\mu))^*$; in Restle's (1970) notation, would be encoded as $N^*(1)$.

Here $S5 = S0$, $S4 = S1$, $S3 = S2$, or, in general, $Si = S(5 - i)$; $i(0{:}\ 5)$. Defining $ck(i)$ as the k's complement of i, we can write instead: $Si = S(c5(i))$.

Restle (1970) introduces a relation which he calls m (for "mirror") in order to permit complementation. He does not, however, use the relation consistently, and hence his various examples do not all fit the same definition of it.[6]

In the example above, the complementation relation is applied to the postscripts designating the locations of symbols. If the sequence uses the alphabet of integers, the relation can also be applied to the symbols themselves. That is to say, Sequence 5 can also be described by $Si = c4(S(i - 3))$; $i(3{:}\ 5)$; $Sj = (j + 1)$; $j(0{:}2)$.

In summary, all of the pattern languages under consideration describe sequences with the help of the relations s, n, p, $+$, $-$, and ck, or some subset of these. The latter three relations are defined only on integers and are in any case definable in terms of n and p. The first three relations are defined on pairs of symbols drawn from any alphabet (ordered set); hence, they are common structural elements for virtually all kinds of linear patterns.

Designation of Symbols

While all of the pattern languages use essentially the same small set of relations, they differ widely in the notation they employ to name or designate the symbols that enter into the relations. There appear to be three different ways to designate symbols: (*a*) naming them explicitly by their positions in the sequence (e.g., $S4$ is the symbol that follows $S3$ and precedes $S5$), (*b*) displaying relations in the pattern in the order in which the corresponding symbols appear in the original sequence, so that the pattern becomes a "template" for the sequence,[7] and (*c*) notating successive runs of identical symbols by numbers corresponding to the run lengths.[8] In the examples thus far, we have used the first method, naming each symbol explicitly by postscripting "S" with its order number in the sequence (numbering from zero). This is a very flexible notation, but it is not compact. Consider the sequence:

$$\text{ABMCDM}\cdots. \qquad [6]$$

In this notation the sequence would be described by the relations:

$$\begin{aligned} S0 = \text{A};\ S3i &= nS(3i - 2); \\ S(3i - 2) &= n(S3(i - 1)); \qquad [7] \\ S(3i - 1) &= \text{M};\ i(1{:}^*) \end{aligned}$$

The postscripts do not reflect very clearly the periodic or hierarchic structure of Sequence 6, in which each triad of symbols repeats the relations of the previous triad. The periodicity becomes more visible if each symbol is designated by a double postscript, the first component denoting the period, and the second, the location (0, 1, or 2) within the period. With this scheme, the second symbol in the sixth period would be designated $(S(5,1))$. The pattern can now be rewritten as:

$$\begin{aligned} S0 = A;\ S(i + 1, 0) &= nS(i,1); \qquad [8] \\ S(i,1) &= nS(i,0); \\ S(i,2) &= \text{M};\ i(0{:}^*) \end{aligned}$$

This notation, employed by Simon and Sumner (1968), still does not fully reveal the structure of Sequence 6, which consists of a simple alphabetic sequence (the first two symbols of each period) interwoven with a repetition of the letter M. The notation can be further simplified, using the second method of designating symbols, by writing down the relations in a definite order, and by introducing names for *working memories*.

---

[6] The example in his Figure 2, as well as the first two examples in his Table 1, is inconsistent with his definition of the relation m.

[7] More formally, this scheme retains some of the information about the order of symbols in the sequence by an order-preserving many–one mapping of symbols of the sequence onto subsequences of the pattern. The number of symbols mapped onto successive sub-sequences of the pattern defines the *period* of the sequence.

[8] As we shall see, this scheme is a specialization and abstraction of the preceding one that eliminates redundant symbols from the pattern, retaining only the superscripts that denote numbers of iterations.

A working memory, call it μ, contains the name of an alphabet and a pointer to a particular location of the alphabet. Moving the pointer locates the *next* location on the alphabet. Using this idea we can notate the pattern of Sequence 6 by:

$$[\mu \leftarrow \alpha](n(\mu)n(\mu)\mathrm{M})^* \qquad [9]$$

which can be read: "Set the pointer in μ to the name of the Roman alphabet, α. Move the pointer one place down the alphabet (i.e., to $n(\mu)=\mathrm{A}$), then one more place (i.e., to $n(\mu) = \mathrm{B}$); then produce the letter M; repeat these three processes indefinitely." (The square brackets in Pattern 9 denote an operation that does not produce a symbol.)

This new notation abandons entirely the use of postscripts to designate symbol locations. Instead, the rules for the three positions in the basic triad are written down, one following the other, in order.

As a result of these two changes in notation—introducing working memories and dispensing with the explicit names for symbols at designated locations—the description is greatly abbreviated. In place of the 52 symbols in Pattern 7, or the 51 in Pattern 8, Pattern 9 contains only 17 symbols. The notation of Pattern 9 is essentially that introduced in Simon and Kotovsky (1963).

Restle (1970) later introduced a different device to permit the suppression of explicit names for symbols. He employed operations upon *sequences* of symbols, the effect of such an operation being to change all of the symbols in the sequence in exactly the same way. To see the relation between his notation and that of Simon and Kotovsky, we denote by S_i a sequence of symbols, $S1$, $S2$, ... Then we can define, for any operator, o,

$$o\{S_i\} = \{o(Si)\} \qquad [10]$$

so that, for example,

$$n\{\mathrm{A,B,C}\} = \{n(\mathrm{A}),\ n(\mathrm{B}),\ n(\mathrm{C})\} = \{\mathrm{B,C,D}\}.$$

Restle also allows concatenation of operators:

$$(o1,\ o2)\{\mathrm{S}_i\} = o1\{S_i\}o2\{S_i\} \qquad [11]$$

that is, the effect of applying the concatenated operator, $(o1,o2)$, to a sequence $\{S_i\}$ is to produce the sequence $o1\{S_i\}$ followed by the sequence $o2\{S_i\}$.

Restle introduces one other useful notational device. Let a lowercase letter, o, designate any operator; then the corresponding capital letter, O, designates the concatenation of s with o. That is,

$$O = s,o \qquad [12]$$

so that $O\{S_i\} = s\{S_i\}o\{S_i\} = \{S_i\}\{o(S_i)\}$. While Restle's notational innovations are applicable only to very special classes of patterned sequences, they permit sequences belonging to these classes to be described perspicuously and compactly. Consider, for example,

$$\mathrm{DEFGEFGHFGHI} \cdots. \qquad [13]$$

In Restle's notation, this can be described by:

$$N^*(N^3(\mathrm{D})) \qquad [14]$$

where

$$N^i\{S_i\} = \{S_i\}N^{i-1}\{n(S_i)\};$$
$$N\{S_i\} = \{S_i\}\{n(S_i)\},$$

whence:

$$N^*(N^3(\mathrm{D}))\rightarrow N^*(\mathrm{D}N^2(\mathrm{E}))\rightarrow N^*(\mathrm{DE}N(\mathrm{F}))$$
$$\rightarrow N^*(\mathrm{DEF}\,n(\mathrm{F}))\rightarrow N^*(\mathrm{DEFG})$$
$$\rightarrow[\mathrm{DEFG}\,N^*(\mathrm{EFGH})]$$
$$\rightarrow[\mathrm{DEFGEFGH}\,N^*(\mathrm{FGHI})]\rightarrow \text{etc.}$$

In the Simon and Kotovsky (1963) notation, the description of Sequence 13 might take the form:

$$[\mu_2\leftarrow\mathrm{D}]([\mu_1\leftarrow\mu_2]\mu_1[n(\mu_1)]^3[n(\mu_2)])^* \qquad [15]$$

Pattern 14 requires only nine symbols, while Pattern 15 requires 27 to describe Sequence 13.

On the other hand, many sequences cannot be described at all in Restle's special notation—Sequence 6 being a simple example. The reason is easy to see. In Sequence 6, unlike Sequence 13, the symbols of the *i*th period cannot be obtained by applying the same operator to the corresponding symbols of the previous period.

The notation used by Leeuwenberg (1969) resembles most closely that of Simon and Kotovsky (1963). In fact, when the 15 patterns of the Thurstone Letter Series Completion Test are encoded in both notations, the pattern length in the Simon and Kotovsky notation has a .8 correlation with the pattern length in Leeuwenberg's notation. Leeuwenberg used a rather elaborate system of punctuation to obviate the need for naming working memories explicitly. Essentially, his idea in encoding a pattern like Sequence 6 is to describe separately each of the intertwined subpatterns, then to indicate which of the symbols in each period belongs to each of the subpatterns. Sequence 6, for example, is constructed from two subpatterns, one providing the first two symbols in each period; the other, the third symbol. (Details of Leeuwenberg's encoding, which involves the notation of "runs" of the subpatterns [see below] can be found in his paper.)

In the schemes of Glanzer and Clark (1962) and Vitz and Todd (1969), matters become simpler because of the restriction to binary (or ternary) alphabets and to the relation *s*. In the binary case, a periodic pattern can be described in terms of the successive run lengths of 1s and 0s, respectively. Thus, the sequence

00100011111··· [16]

can be encoded as (02135)*. (By convention the sequence starts with a run of 1s, in this case, of length zero.) The longer the runs, on average, the more efficient is this encoding. The least parsimonious description—that for a simple alternation—is simply the sequence itself.

With a ternary alphabet, each run can be described by juxtaposition of the symbol and run length. Thus,

AAABBBCCC··· [17]

can be described as (A3B3C3)*.

In terms of our previous notations, we could describe Sequence 16 by:

$$((s(0))^2\ (s(1))^1\ (s(0))^3\ (s(1))^5)^* \quad [18]$$

Thus the run-length notation can be derived from the more general one by omitting all of the redundant, predictable symbols in the explicit notation and retaining only the superscripts.

Still more complicated codes may facilitate the description of special classes of binary patterns. If patterns consist, for example, of runs *and* simple alternations on a binary alphabet, then it may be efficient to recode into a ternary alphabet—$1 \rightarrow \alpha$, $0 \rightarrow \beta$, $01 \rightarrow \gamma$—and then use the ternary code suggested above for Sequence 17 (cf. Feldman, 1961).

Where one symbol occurs much more frequently than the other in a binary sequence, a more efficient coding may be achieved by simply mentioning those positions in the sequence that are occupied by the rarer symbol, for example $S1 = S12 = S19 = 1$. By convention, any symbol not explicitly mentioned is the common one (in the example, 0). This method of abbreviation is a special case of a more general Principle of Exceptions. Normal values are defined for certain symbols, and the symbol is not explicitly mentioned when it takes on its normal value. Thus, in the Simon and Kotovsky (1963) notation, we can take *s* as the normal value of the operator on a working memory, and simply write μ in place of $s(\mu)$ wherever the latter occurs.

Basically, all of the representations we are considering describe sequences by defining the relations that determine symbols in terms of preceding symbols. The length of code for any given sequence (hence the complexity of the pattern) depends basically on the number of relations that must be defined. However, some coding economies can be realized for special classes of sequences by using the devices just mentioned; for example, by omitting mention of common elements, or by exhibiting explicitly only nonnormal values of symbols.

These devices are used by human subjects to achieve succinct descriptions of sequences. Glanzer and Clark (1962), whose method involved asking subjects to report patterns verbally, found that the subjects generally described binary patterns either in terms of runs (i.e., in a language like that used above to describe Sequence 16, or by naming the location of the "exceptional" symbols (e.g., 1s in positions 2 and 6"—a simplification, for binary sequences, of the Simon & Sumner, 1968, language).

Vitz and Todd (1969), as previously noted, constructed an information-theoretic measure of pattern complexity. The encoding system underlying their measure was based on the detection in sequences of runs of individual symbols and runs of subpatterns (see Vitz & Todd, 1969, pp. 435–436). As a result, there is a very high correlation between their information measure and a simple count of runs. Table 1 shows, for four sets of data presented by Vitz and Todd, the correlation coefficients (*a*) between Vitz and Todd's information measure (H) and sequence complexity (C) as judged by subjects; (*b*) between H and length of code (L), based on runs; and (*c*) between C and L.

We see from Table 1 that H accounts for 80% to 95% of the variance in judged C of the sequences. Except in one case, where the correlation is .79, the correlation of H with L is around .9. Hence variation in code length accounts for about 80% of the variance in amount of information contained in these sequences. If H were the "true" estimator of C, and L an approximate surrogate for H, otherwise unrelated to C, then the correlation between L and C would be $r'_{\mathrm{LC}} = r_{\mathrm{HC}} \times r_{\mathrm{HL}}$. As Table 1 shows, the actual correlation, r_{LC}, is always larger than r'_{LC}. Hence code length makes a (small) independent contribution to the prediction of complexity, over and above the part explanable by the relation of L to H.

TABLE 1

CORRELATION OF CODE LENGTH, INFORMATION CONTENT, AND JUDGED COMPLEXITY OF FOUR SETS OF PATTERNED SEQUENCES

Variable		Set of sequences			
		1	2	3	4
1. H_{VT} × complexity	r_{HC}	.98	.95	.94	.91
2. H_{VT} × code length	r_{HL}	.92	.79	.91	.88
3. Code length and complexity	r_{LC}	.91	.79	.89	.90
4.	r'_{LC}	.90	.75	.85	.80

Note.—H_{VT} = Vitz-Todd information measure; code length = number of runs of identical symbols; complexity = judged complexity (Sets 1–3), or number of reproduction errors (Set 4). Data for the four sequence sets from Vitz and Todd (1969), Tables 6, 7, 8, and 9, respectively. r'_{LC} = predicted correlation between code length and complexity. $r'_{\mathrm{LC}} = r_{\mathrm{HC}} \times r_{\mathrm{HL}}$.

In summary, the pattern languages we are examining use a variety of notations to designate the symbols between which specified relations hold. There is a tradeoff, in general, between the generality of a notation—the range of pattern types it can describe—and its efficiency in describing particular types of patterns. Of the notations reviewed here, that of Simon and Sumner (1968) is the most explicit and general, but also the most "verbose" in designating locations. The Simon and Kotovsky (1963) notation is considerably more concise, giving up only a little generality. The notation of Restle (1970) is quite specialized but parsimonious where applicable. The coding schemes considered by Glanzer and Clark (1962) and Vitz and Todd (1969) are even more specialized and concise. They are obtained by omitting all the information that is redundant for the special sequences to which they are applicable.

Iteration and Hierarchy

Most of the coding schemes use numerical superscripts or subscripts to indicate repetition of pattern elements—operators, symbols, or subpatterns. Thus n^2 is a convenient notation for "next of next," so that $n^2(\mathrm{A}) = n(n(\mathrm{A})) = \mathrm{C}$. Similarly, $[\mu \leftarrow \mathrm{A}](n(\mu))^3$ provides a shorthand for $[\mu \leftarrow \mathrm{A}](n(\mu)n(\mu)n(\mu)) = \mathrm{BCD}$. We have seen how a run-length code can be derived by abstraction from such a notation for iteration:

$$11100111100 \rightarrow (1)^3(0)^2(1)^4(0)^2 \rightarrow 3242.$$

The structure of patterns is not limited to simple periodicity, but may involve a hierarchy of periods. Restle (1970) has paid particular attention to hierarchic patterns, which were earlier used to a

TABLE 2

RESULTS OF OPERATION

Operations	Push-down list	Symbols produced	Restle's operators
$[\mu \leftarrow 1]$	1		1
A push	11		
B push	111		
$\mu\, n(\mu)$	211	12	N
Pop	11		
C $(\ )^2$	(Repeat from B to C)	12	S
$[n(\mu)]$	21		
D $(\ )^2$	(Repeat from B to D)	23 23	N
Pop	1		
E $(\ )^2$	(Repeat from A to E)	12 12 23 23	S
$[C7(\mu); n \leftarrow p]$	6		
F $(\ )^2$	(Repeat from A to F)	$(65\ 65\ 54\ 54)^2$	C7

limited extent by Kotovsky and Simon,[9] and extensively by Simon and Sumner (1968) in their description of music.

Consider Restle's (1970) example:

$$121223231212232365655454656554 \quad [19]$$

The hierarchic phrase structure of this sequence can be made visible by punctuation:

$$(12\ 12)\ (23\ 23),\ (12\ 12)\ (23\ 23);$$
$$(65\ 65)\ (54\ 54),\ (65\ 65)\ (54\ 54), \quad [20]$$

With our adaptation of Restle's notation, we can describe the sequence succinctly, thus:

$$C7(S(N(S(N(1))))), \quad [21]$$

and from this pattern we can regenerate the sequence again by successive steps:

$$\begin{aligned} &C7(S(N(S(N(1)))))\rightarrow C7(S(N(S(12))))\\ &\rightarrow C7(S(N(12\ 12)))\rightarrow C7(S(12\ 12\ 23\ 23))\\ &\rightarrow C7(12\ 12\ 23\ 23\ 12\ 12\ 23\ 23)\\ &\rightarrow (12\ 12\ 23\ 23\ 12\ 12\ 23\ 23\\ &\qquad 65\ 65\ 54\ 54\ 65\ 65\ 54\ 54) \quad [22] \end{aligned}$$

There is a serious objection to taking Pattern 21 as a hypothesis of how a subject might encode Sequence 20. As Sequence 22 shows, regenerating the sequence from the pattern calls for a short-term memory capacity of at least 16 symbols of the sequence as input.

An alternative encoding, based on the Simon and Kotovsky (1963) notation, avoids this difficulty. To explain it, it is necessary to introduce the concept of *push-down list*. A push-down list is simply a sequence that can be altered only by adding a symbol to, or taking a symbol from, its front end. Let us call *push* the operation of duplicating the first symbol of a push-down list, and *pop* the operation of removing the first symbol. Then push(ABC) = AABC, while pop(ABC) = BC.[10] We can use a working memory, μ, holding a push-down list, and these operations to encode Sequence 19 as:

$$[\mu \leftarrow 1]((\text{push }((\text{push }\mu\, n(\mu)\text{ pop})^2[n(\mu)])^2\ \text{pop})^2[C7(\mu);\, n \leftarrow p])^2 \quad [23]$$

where $[n \leftarrow p]$ replaces Operation n by Operation p everywhere throughout the pattern.

Let us follow step by step (see Table 2) the generation of the sequence from Pattern 23, indicating also the correspondence between these steps and Restle's operators in Pattern 21.

Now in this process, unlike Restle's, only the symbols in the push-down list need be held in short-term memory (a maximum of three symbols), not all of the symbols of the sequence already produced (a maximum of 16 symbols). From this standpoint, the representation employing the push-down list provides a more plausible model of the psychological process than the recursive representation. The same argument applied to Patterns 14 and 15 for generating Sequence 13 shows that the former pattern (Restle, 1970) calls for a maximum of four symbols in short-term memory, the latter (Simon & Kotovsky, 1963), a maximum of two. With the use of a push-down list, Pattern 15 could be rewritten:

$$[\mu \leftarrow D]\ (\text{push }\mu\ (n(\mu))^3\ \text{pop }[n(\mu)]^* \quad [25]$$

---

[9] K. Kotovsky and H. A. Simon. Empirical tests of a theory of human acquisition of concepts for sequential patterns. In preparation.

[10] Push-down lists were invented along with computer list-processing languages in the mid-1950s. They have been used widely in implementing parsing programs (so-called "push-down automata") for computing languages and natural languages. Hence, the notion that a push-down list may be part of the organization of short-term memory already has some currency and support in psycholinguistics (see Yngve, 1960).

In Restle's (1970) recursive scheme, the short-term memory requirements depend on the length of the period at the highest level of the pattern hierarchy. This number can be expected to increase (approximately) geometrically with number of levels. In the scheme using a push-down list, the short-term memory requirements increase only linearly with number of levels in the hierarchy.

Restle has produced some data that demonstrate quite convincingly (e.g., Restle, 1970, Figure 3) the psychological reality of the hierarchic phrase structure of patterns. However, his data are equally compatible with the encodings represented by Patterns 21 and 23, for both postulate the same phrase structure in the pattern.

In summary, various of the encoding schemes make provisions for iteration of subpatterns and for hierarchic phrase structure of patterns. Empirical evidence can be (and has been) produced to demonstrate that subjects, in fact, encode some patterns hierarchically. The different encoding schemes place quite different demands on short-term memory, however, in generating sequences from their patterns.

ADDITIONAL PSYCHOLOGICAL IMPLICATIONS

The analysis up to this point has been concerned mainly with describing different languages that have been proposed for encoding patterned sequences. We have introduced behavioral evidence only occasionally to elucidate some particular point of comparison or contrast. The remainder of this paper is concerned with additional implications of coding languages for psychological theory.

If a particular sequence can be encoded, and represented internally, in several different ways, then it is important to explain what determines which encoding a human subject will use. If different subjects use different encodings, it is important to explain the causes for the difference, and its consequences in terms of task performance.

The tasks that subjects have performed in the laboratory with patterned sequences include: (*a*) ranking sequences according to complexity, (*b*) inducing patterns from sequences, (*c*) describing sequences verbally, (*d*) learning and remembering sequences, and (*e*) extrapolating sequences.

Task *a* only produces observable behavior when combined with *c*, *d*, or *e*. Any one of the latter can be used, in turn, to produce measures of sequence "difficulty"—that is, difficulty can be measured by the time required to perform any one of these tasks with a given sequence, or by number of errors. (Not only the number of errors, but also the precise nature of the errors can be determined.)

A priori measures of sequence complexity can also be constructed and these can be correlated with sequence difficulty as measured by subjects' behavior. As we have seen, sequence complexity may be measured by code length or by information content; the two kinds of measures are mutually translatable, the one into the other; and both kinds of measures are defined relative to a specific encoding scheme.

Correlation of A Priori Complexity Measures

When two or more a priori complexity measures are computed for the same set of sequences, the correlation among these will indicate the similarities of the encodings or, conversely, the sensitivity of the measures to details of encoding.

We have already seen that Leeuwenberg's (1969) measure correlates .8 with Simon and Kotovsky's (1963) when both are applied to a set of 15 sequences from the Thurstone Letter Series Test, and that this high correlation is a consequence of the basic similarities in the coding concepts underlying both schemes.

Similarly, a simple computation verifies that Vitz and Todd's (1969) information measure correlates at about the .9 level with code length when binary and ternary sequences are encoded in terms of simple run lengths.

Since we emphasized at the outset that code length is relative to the coding language, it is not obvious why these high correlations occur. They occur because the various coding schemes are not unrelated but are all based on the same set of relations (n, p, s, c), and on notation for iteration. When we compare Patterns 14 and 15 for Sequence 13, for example, we see that, however more compact Pattern 14 is, each symbol in that pattern corresponds to a symbol, or set of symbols, in Pattern 15. Thus, we have the correspondences:

$$\begin{aligned} &N^* \leftrightarrow ([\mu_1 \leftarrow \mu_2]\cdots[n(\mu_2)])^* \\ &N^3 \leftrightarrow \mu_1(n(\mu_1))^3 \\ &D \ \leftrightarrow [\mu_2 \leftarrow D]. \end{aligned}$$

As long as we do not change the basic relations used, we can expect that *relative* complexity of sequences will be rather insensitive to details of the coding, providing that the sequences under consideration are not too heterogeneous. Hence, our ability to predict measures of behavior may not depend on our knowing in detail the exact encoding scheme that subjects use. If this is so, it will be difficult to choose alternative encoding hypotheses without sophisticated analysis of the behavioral data.

Correlation of A Priori with Judged Complexity

A major use of complexity measures (see e.g., Vitz & Todd, 1969) has been to predict complexity as judged by subjects from an a priori measure of complexity. For the reasons just stated, the achievement of high correlations cannot be taken as a validation of the specific a priori measure used. Rather, the high correlations should be interpreted to imply: (a) that subjects confronted with sequences do in fact encode them with the aid of the relations we have discussed, and (b) that the length of the encoded pattern is the basis for the subjects' judging the complexity of the sequence.

This general hypothesis was put forth independently by Glanzer and Clark (1962) and by Simon and Kotovsky (1963). Glanzer and Clark went further, however, and, in their "verbal loop hypothesis," asserted that the encoding was verbal (i.e., consisted of a sequence of English words and phrases). They said, specifically:

> [The subject] carrying out a perceptual recall task puts the information through a verbal loop. (1) He translates the visual information into a series of words. (2) He holds the verbalization and makes his final response on the basis of that. . . . Complexity, under this hypothesis, becomes identified with the length of the S's [subject's] verbalization [p. 295].

There is nothing in the data, however, to demonstrate that the internal representation is verbal. The verbal descriptions that Glanzer and Clark elicited from their subjects could equally well have been a fairly direct recoding into verbal form of the internally represented symbolic (but nonverbal) structures. It seems preferable, therefore, to adopt the more neutral stance of Simon and Kotovsky (1963) regarding the nature of the internal representation:

> Subjects attain a serial pattern concept by generating and fixating a pattern description of that concept. . . . [Subjects] have stored in memory a program capable of interpreting and executing descriptions of serial patterns [p. 538].

Complexity and Difficulty

A second use of complexity measures, whether a priori or judged complexity, is to predict the relative difficulty subjects will experience in performing specific tasks on different sequences.

It is not obvious that the ranking of sequences by difficulty will be the same for all tasks—the same for describing sequences verbally as for learning periodic sequences as for extrapolating nonperiodic sequences. To describe a sequence verbally, the subject must (a) discover the pattern and (b) hold the pattern in memory (short term or long term) while he produces it (Glanzer & Clark, 1962). To learn a periodic sequence he must either (a) fixate the sequence by rote or (b) discover the pattern, hold it in memory, and use it to generate the sequence (Gregg, 1967). To extrapolate a nonperiodic sequence he must (a) discover the pattern, (b) remember it, and (c) use it to generate the sequence (Simon & Kotovsky, 1963).

An EPAM-like learning theory would suggest that three or four different sources of task difficulty are present in different combinations in these tasks: (a) difficulty of *discovering* a pattern, (*b*) difficulty of *fixating* either a sequence or a pattern in *long-term memory*, and (*c*) difficulty of holding in *short-term working memory* the place keepers or pointers required to produce the sequence from the pattern.

However, disentangling these components of difficulty empirically is not at all easy. In general, the more complex the pattern (as measured by code length), the more relations it is necessary to induce from the sequence in order to find the pattern, the more symbols it is necessary to fixate in order to store the pattern in long-term memory, and the more working memories are necessary in order to produce the sequence from the pattern. Because of this confounding of components of difficulty, all investigators have found code length to be a reasonably good predictor of difficulty, in whichever way difficulty is measured. Thus, Glanzer and Clark (1962) report a correlation of −.826 between mean verbalization length (code length) and accuracy of recall; Vitz and Todd (1969), using the same data, find a correlation of −.82 between information content (which is interpreted here as an alternative estimate of code length) and accuracy of recall. Leeuwenberg's (1969) data show correlations of .65 and .66 between two measures of difficulty (errors and time, respectively, for a group of 12 subjects) on the Thurstone Letter Series Task and code length, using his encoding scheme; there are similar correlations, of .74 and .61, between these same two measures of difficulty and code length with the Simon and Kotovsky (1963) coding scheme.

Direct Evidence of Encoding

Three kinds of "direct" evidence have been sought to determine the actual encodings used by subjects. Glanzer and Clark (1962) simply instructed the subject to describe each sequence, with the results already mentioned. Kotovsky (1970) and Kotovsky and Simon (see Footnote 9) obtained verbal protocols and other evidences of subjects' sequential behaviors while discovering the patterns in a set of sequences. Williams (1969) obtained a combination of verbal protocols and eye-movement data. The patterns inferred from this direct evidence generally fit well the Kotovsky and Simon coding language. While numerous individual differences were found, these almost all fell within the general range of concepts allowed by the formal language—that is, it was usually possible to express the variant encodings in that language.

For example, in the specific case of Sequence 13, some eight of Kotovsky and Simon's 14 subjects used a pattern description corresponding closely to Pattern 15 or 23, and only one a description resembling Pattern 14. On the other hand, with the slightly simpler pattern, PON ONM NML . . . , four subjects used an encoding resembling Restle's (1970). Restle's description only requires three working memories to handle the latter pattern, but four to handle Pattern 14, and this difference in the demands upon working memory may account, in whole or part, for the preferences among encodings.

Gregg (1967) has also provided direct evidence on how subjects describe a pattern they are trying to discover and, more important yet, how their mode of description affects the ease or difficulty of learning the pattern.

CONCLUSION

A survey of the principal alternative formulations of a theory to explain human performance in tasks involving patterned sequences shows all of these formulations to be mild variants on a basic theme. The formulations agree in proposing the following: (*a*) that subjects perform these tasks by inducing pattern descriptions from the sequences; and (*b*) that these pattern descriptions incorporate the relations of *same* and *next* (on familiar alphabets) between symbols, iteration of subpatterns, and hierarchic phrase structure.

As a result of this concensus on fundamentals, measures of the relative complexity of different patterns can be expected to be nearly the same, independently of which specific encoding languages are used in computing them. A little direct evidence is now available to help us understand what circumstances subjects will use one or another of these encodings (or others).

It is quixotic to seek crucial experiments to determine *the* coding language, because (*a*) different subjects almost surely use different representations for the same sequences, and (*b*) the representations used may well be modified by task instructions, previous experience, and the character of the set of sequences used in the task. Empirical evidence on these points has already been mentioned. The same conclusion is suggested by the fact that the correlations between subjects in ratings of complexity are far from perfect. (In their paper, Vitz & Todd, 1969, mention mean or median among-subject correlations of .67 and .84 for two sets of data.) Instead of seeking to discover "the" pattern code or "the" complexity measures, we need to determine the conditions (sociological and psychological) under which subjects will adopt one or another coding scheme in handling patterned sequences.

REFERENCES

ESTES, W. K., & BURKE, C. J. Application of a statistical model to simple discrimination learning in human subjects. *Journal of Experimental Psychology,* 1955, **50**, 81–88.

FELDMAN, J. An analysis of predictive behavior in a two-choice situation. Unpublished doctoral dissertation, Carnegie-Mellon University, 1959.

FELDMAN, J. Simulation of behavior in the binary choice experiment. *Proceedings of the Western Joint Computer Conference,* 1961, **19**, 133–144.

FELDMAN, J., TONGE, F. M., JR., & KANTER, H. Empirical explorations of a hypothesis-testing model of binary choice behavior. In A. C. Hoggatt & F. E. Balderston (Eds.), *Symposium on simulation models.* Cincinnati, Ohio: South-Western Publishing, 1963.

GLANZER, M. S., & CLARK, H. H. Accuracy of perceptual recall: An analysis of organization. *Journal of Verbal Learning and Verbal Behavior,* 1962, **1**, 289–299.

GOODNOW, J. J., & PETTIGREW, R. F. Effect of prior patterns of experience upon strategies and learning sets. *Journal of Experimental Psychology,* 1955, **49**, 381–398.

GREGG, L. W. Internal representation of sequential concepts. In B. Kleinmuntz (Ed.), *Concepts and the structure of memory.* New York: Wiley, 1967.

KLAHR, D., & WALLACE, J. G. The development of serial completion strategies; an information processing analysis. *British Journal of Psychology,* 1970, **61**, 243–257.

KOTOVSKY, K. An empirical test of the Simon and Kotovsky "concept former" model of human letter series sequence extrapolation behavior. Unpublished master's thesis, Carnegie-Mellon University, 1970.

LAUGHERY, K. R., & GREGG, L. W. Simulation of human problem-solving behavior. *Psychometrika,* 1962, **27**, 265–282.

LEEUWENBERG, E. L. L. Quantitative specification of information in sequential patterns. *Psychological Review,* 1969, **76**, 216–220.

PIVAR, M., & FINKELSTEIN, M. Automation, using LISP, of inductive inference on sequences. In E. C. Berkeley & D. G. Bobrow (Eds.), *The programming language LISP.* Cambridge, Mass.: Information International, 1964.

RESTLE, F. Grammatical analysis of the prediction of binary events. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6**, 17–25.

RESTLE, F. Theory of serial pattern learning: Structural trees. *Psychological Review,* 1970, **77**, 481–495.

SIMON, H. A., & FEIGENBAUM, E. A. Elementary perceiver and memorizer: Review of experiments. In A. C. Hoggalt & F. E. Balderston (Eds.), *Symposium on simulation models.* Cincinnati, Ohio: South-Western Publishing, 1963.

SIMON, H. A., & KOTOVSKY, K. Human acquisition of concepts for sequential patterns. *Psychological Review,* 1963, **70**, 534–546.

SIMON, H. A., & SUMNER, R. K. Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgment.* New York: Wiley, 1968.

VITZ, P. C., & TODD, R. C. A coded element model of the perceptual processing of sequential stimuli. *Psychological Review,* 1969, **76**, 433–449.

WILLIAMS, D. S. Computer program organization induced from problem examples. In H. A. Simon & L. Siklóssy (Eds.), *Representation and meaning.* Englewood Cliffs, N. J.: Prentice-Hall, 1972.

WILLIAMS, D. S. Computer program organization induced from problem examples. Unpublished doctoral dissertation, Carnegie-Mellon University, 1969.

YNGVE, V. A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society,* 1960, **104**, 444–446.

(Received April 24, 1972)

Psychological Review
1972, Vol. 79, No. 5, 383–393

DUAL FUNCTIONS OF CONSCIOUSNESS

TIM SHALLICE[1]

University College London, England

It is argued that theoretical developments in cognitive psychology and the increasing use of introspective reports require a rationale, and that this should involve consideration of consciousness. An information-processing model is suggested for cognitive processes, based on the ideas of G. A. Miller, E. Galanter, and K. H. Pribram. Its main characteristic is that only one "action system" (similar to a plan) can be maximally activated at any time, but that other such systems may be more weakly activated. The model is derived by cybernetic analogies, and empirical evidence for it is briefly reviewed. It is suggested that part of the model—the selector input to the dominant action system—has properties which correspond to those of consciousness. The selector input has the functions of selecting which action system is dominant and of setting its goal.

In this paper it is argued that the basic phenomenological concept—consciousness—can be mapped onto an information-processing concept. In arguing for this mapping, the paper has two aims. The major aim is to remove an anomaly in information-processing theory, namely, that while many aspects of mental life have been modeled, theorists have avoided direct consideration of conscious awareness, although this is vital if cognitive psychology is to have a satisfactory theoretical base. The minor aim is to facilitate the interpretation of experiments which involve the subject in complex introspections, judgments, or strategies.

In the development of models, cognitive psychologists have approached the problem of consciousness indirectly as in work on attention (e.g., Deutsch & Deutsch, 1963; Posner & Boies, 1971) or have used it as an unexplained component as, say, in Waugh and Norman's (1965) requirement that transfer from primary to secondary memory requires conscious rehearsal. There has been no explanation of how such a phenomenological concept can be related to ideas about mechanism. One possibility is to reject the need for such an explanation, to argue that consciousness is a pseudoconcept. Yet, increasingly, concepts such as strategy and rehearsal are being used as explanatory concepts. Such concepts depend on the theorist reflecting on his conscious experience and are very different from a concept like primary memory whose roots lie in physiology and computer technology. Even a concept like eidetic imagery is not understood by a person lacking it as he would a mechanistic concept, but through descriptions by other people. Hence, if consciousness is to be rejected as a pseudoconcept, to be consistent such concepts must be rejected too.

This would much restrict the present explanatory repertoire, but it would also have other repercussions. Subjective reports are increasingly used, if on occasion slightly disguised under the label "protocols." Also more use is being made of instructions which require subjects to monitor their conscious experience in complex ways as for Sperling's (1967) subjects who had to adjust the time of occurrence of a click so that it coincided subjectively with the termination of a visual image. If consciousness is a concept as unrelated to reality as a concept like "ghost," how such experiments produce scientifically useful results becomes very mysterious and unexplained. It thus seems that the problem of consciousness occupies an analogous position for cognitive psychology as the problem of language behavior does for behaviorism, namely, an unsolved anomaly within the domain of the approach.

---

[1] I should like to thank David Gaffan, Ted Honderich, A. R. Jonckheere, Phil Johnson-Laird, Ulric Neisser, and Peter Wason for their most helpful comments on earlier drafts of this manuscript.

Requests for reprints should be sent to Tim Shallice, Psychology Department, National Hospital, Queen Square, London, WC1, England.

Turning to the second aim, phenomenological evidence remained acceptable, at least in perception, even when psychology was at its most behaviorist. Judgments required of subjects would be of standard sorts, for example, psychophysical ones, and the theoretical explanations given for the results could normally be in terms of mechanisms clearly distinct from the judgment process itself. For instance, the results of color-matching experiments are interpretable in terms of retinal mechanisms, while the matching process itself is central. Moreover, the assets and pitfalls of particular judgmental procedures became known (see Woodworth & Schlosberg, 1954), alternative judgment procedures could be used to cross-check inferences, and potential judgmental artifacts can be investigated by hypothesis testing—the converging operations of Garner, Hake, and Eriksen (1956). In effect, the judgmental procedures become almost as objective a source of data as a physical measuring instrument such as an electrode (see Attneave, 1962, for a related view).

However the more central the mechanism under investigation, or to use Miller, Galanter, and Pribram's (1960) term, the less it belongs in the Image, the more difficult do such cross-checking operations become. The introspection or judgment process will be intimately involved with the mechanisms under investigation so the two stages will not be separable as measuring instrument and phenomenon. As an illustration, consider the difficulty of demonstrating that De Groot's (1965) protocols of chess thinking are free of artifacts. His main results depend on obtaining verbalizations from chess masters, as they analyze difficult positions, a procedure that cannot be simulated at present and only in a very general way, repeated. His findings provide some of the best support for the primarily serial-processing nature of complex thought, but they depend on the uncritical acceptance of introspective evidence—one of the best learned dangers in psychology.

One means of avoiding this danger is to develop theories of the introspective process so that one can assess theoretically what should or should not be subjectively reportable. This assumes that the introspective process can be split into two parts—the production of conscious experience and the reporting of the experience. The primary evidence for making this split is the way in which experimenters, for example, in perceptual experiments, utilize their own experience in the experimental situation in order to interpret the subjects' reports (Natsoulas, 1967; Rosner, 1962). The extremely complex problems relating to the second stage have been interestingly discussed by Natsoulas (1967, 1970); they will not be considered in this paper.

Two linked assumptions are made in attacking the problem of the relation between conscious experience and mechanism. First, the identity theory of the relation between mind and brain will be assumed correct. This is that mental processes are identical with certain physical processes in the brain in the sense of "identity," by which a gene is identical with a section of a deoxyribonucleic acid (DNA) molecule. Following from this, the problem of relating information-processing theories to phenomenological generalizations is assumed analogous to that of relating two neighboring fields of science. "Phenomenological generalizations" refers to hypotheses ranging from ones based on simple introspection such as that conscious experience can be of various types, percepts, images, intentions, etc., to more obscure philosophical ones such as that a defining characteristic of a mental state is its property of intentionality.

With respect to the second assumption, it is clear that one technique in science for relating two fields is by constructing isomorphic correspondences between the concepts and operations of one field and those of another. Numerous examples of such constructions exist in the history of science, the most famous recent one being that of molecular biology linking biology and chemistry. As Beckner (1959) showed, using this example, it is not necessary to construct isomorphisms between all the concepts and operations of one paradigm and those of another (i.e., reduction of the first to the second) for the linking to be successful. He argues that the concept of a kidney cannot be reduced to its physical and chemical properties, since it can only be understood in terms of its function within a whole system, and function cannot be related to existing physical and chemical concepts.

This becomes clearer if one uses the analogy of a scientific theory as a map of the world. Two perfectly correct maps of the same area cannot necessarily be totally reduced one to the other, as they may be concerned with different aspects of the area, for example, ancient monuments and geological structure. This approach avoids one claimed difficulty for identity theory (Shaffer, 1961); while conscious experience does not have the property of being spatially localized, brain states do.

In general, it seems that the philosophical objections to the identity theory can be answered (Armstrong, 1968; Borst, 1970). Probably the strongest objections are linked to Wittgenstein's (1953) succinct but opaque claim: "An 'inner process' stands in need of outward criteria [p. 153]." Wittgenstein's approach has frequently been interpreted (e.g., Smart, 1963) to mean that insofar as mental states can be communicated, they can be reduced to observables. It follows that any attempt to use phenomenological information in science which does not just rephrase behavioral information would be vacuous. As pointed out earlier, the practice of current research appears to be proving this claim false. It is likely that Wittgenstein did not hold such a neobehaviorist position (see Malcolm, 1963) and that his arguments are more relevant to the problems surrounding the second stage of the introspective process—the reporting—which are not considered in this paper. His actual argument depends on showing that phenomenological terms do not describe a private world unrelated to action. Therefore, it seems compatible with the present approach in which consciousness is to be related to action.

Proponents of the identity theory have argued for it on the ground of parsimony, given present scientific knowledge. By contrast, Gray (1971) has argued that it offers no guides to what properties of the brain give rise to consciousness, and that it limits the scientific questions one can ask about consciousness. Identity theorists do tend to view consciousness as a property dependent, at least in part, on the nerve cell composition of the brain and not just on its system of organization (e.g., Armstrong, 1968). However the present theory demonstrates that the identity theory can relate to system level properties of the brain and that it can form a basis for theorizing on questions which Gray considered important, such as the survival value of consciousness.

As no presently existing model seems to provide a basis for constructing an isomorphism, one is developed in this paper. It is not conceived as a finished model—other models might well be constructed to account for the limited empirical evidence considered—but the development of a model enables answers to the metalevel questions to be made concrete. Since the argument on the proposed isomorphic relation between phenomenological and information-processing concepts moves over a number of different considerations, a brief summary of the isomorphism is presented here. The brain is conceived of as containing, among other units, a large set of action systems (similar to Miller et al.'s, 1960, TOTE units), only one of which can be maximally activated at any given time. A certain sort of input to each action system has two functions. It determines first whether that action system will become dominant, that is, maximally activated. If so, it then sets the goal of the action system until either the goal is achieved or the action system is reactivated with a different goal. It is this input (the *selector input*) that is identified with the content of consciousness. A flow diagram of the model is shown in Figure 1.

Since direct empirical testing of multiaxiomatic theories is notoriously complex (see Broadbent, 1958), the model is constructed on the basis of cybernetic considerations so as to provide another form of support. It is argued that the model satisfies constraints necessary for any mechanism which would function like a human organism. The cybernetic section is followed by a brief

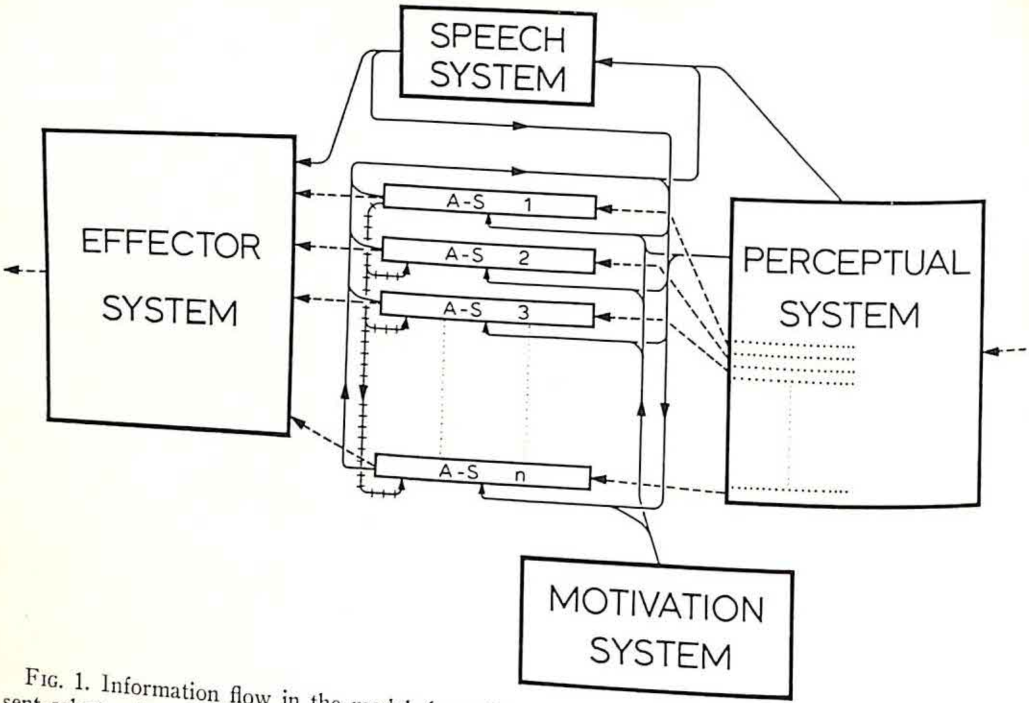


FIG. 1. Information flow in the model (*n* action system [A-S] version). (Solid lines represent selector input. [Speech output via the effector system is included in this category as, on the model, it has the function of providing selector input for other people.] Hatched lines represent inhibitory interactions between action-systems. [To avoid complicating the diagram, the only inhibitory output shown is that from Action System 1.] The dotted lines in the perceptual system represent the partial autonomy required for particular specific input processing. For simplicity, certain interconnections not considered in the models have been left out, e.g., sensory-effector ones. As no arguments have been presented on whether or not a functionally separate memory unit is required, the memory aspects of the model are not represented.)

review of the empirical evidence which is itself followed by more detailed phenomenological considerations.

One claimed characteristic of a conscious state is that it has an object. One is thinking of X, perceiving Y, or intending Z, and only one of these at once. This unitary nature of consciousness is mirrored in psychology by limited-capacity ideas. Evidence, particularly that of Posner and Boies (1971), discussed later, indicates that there is a limited capacity for the programming of motor responses, which suggests developing the model from consideration of the control of actions.

CYBERNETIC CONSIDERATIONS AND MODEL ASSUMPTIONS

A common assumption of theorists is that the infinite variety of one's potential actions is produced by the operation of a finite number of internal control systems (e.g., Piaget's, 1970, schemes; Miller et al.'s, 1960, TOTE units). It has been widely argued that human control systems operate hierarchically in an analogous way to computer programs where the main program may call in routines which may themselves call in other routines and so on (Miller et al., 1960; Newall, Shaw, & Simon, 1958). Miller et al. were primarily concerned with the way particular control systems operated. Instead, in this section, while their general approach is accepted, the problem of how control passes from one system to another is tackled. The unit of control—the action system—is treated as a black box whose properties, stated in later assumptions, are similar to those of programs, except for Assumptions 3 and 4. (The program analogy is not intended to refer to the internal operation of action systems, which are probably most easily conceived in analogue terms following the work on simulation of motor skills; e.g., Bernstein, 1967; Gibbs, 1970; Wiener, 1948.)

Using the program analogy, the basic problem of this section is how a mechanism simulating a human could have the following three properties: (*a*) a very large number of individual programs; (*b*) the routine facility, that is, hierarchical changes of control; and (c) rapid access to and realization of the most vital program. Throughout, the effects of learning are ignored so that only asymptotic performance will strictly be relevant to the model. The assumptions of the model are the following:

1. In addition to perceptual, motivational, and effector systems, the brain contains a set of action systems.
2. Action systems operate so as to attain a goal, by means of output to effector units or to other action systems. Each action system may have a finite number of variables set as a goal when it is initiated.

If more than one action system has a subsystem in common, then it becomes a separate action system or an effector unit. The setting of variables, discussed in Assumption 6, and the separation of common subsystems are both necessary to enable the system to operate in an analogous fashion to a computer program. The separation between action systems and effector units will be dealt with after Assumption 5.

3. The speed and accuracy with which an action system operates depends on its level of activation, which can continuously vary. The minimum level of activation necessary for normal functioning differs between action systems depending on their complexity and the extent to which they have been learned.
4. No more than one action system may be strongly activated (i.e., become dominant) at any given time. This results from the activation of each action system being inhibited by every other action system by an amount which increases monotonically with the other's activation.

The two last assumptions are put forward as a solution to the potential cybernetic problem that an organism has many goals which it needs to achieve at any one time and has only a limited number of effector units available.

It would be possible to arrange that each effector unit is controlled by the most important action system requiring it, by either serial or nonserial means. Serially, a relatively simple program could list the action systems in order of importance. Given the large number of action systems a human would need to possess and the need to make continual changes in priority with the changing stimulus situation, such a system would prove very slow given the relative slowness of neural as opposed to electronic processing. For example, if 2^n action systems existed, to reassess the priority of one action system would require n successive activation comparison operations.

On a simple parallel model operating according to Assumption 3, it would seem possible for each effector unit to be controlled by the action system which activates it most strongly, inputs from other action systems being inhibited. However, consider the situation in which two action systems are roughly equally activated. An example would be changing gear and signaling to go right in an English car. In such a car, changing gear involves movements with the left hand and foot while holding the steering wheel with the right hand; signaling right involves holding the steering wheel with the left hand while using the right arm for signaling. Thus the effector units for both arms receive conflicting input from the two action systems. If the motivating circumstances for changing gear and signaling right are equally strong, then all four inputs to the motor control units will be equally strong (this presupposes that the outputs of an action system to each effector unit are equally strong; other assumptions are also disastrous). Given any noise in the system, there would be a 50% chance of the effector units being under the control of different action systems, if the stronger input to effector units inhibits the weaker. One consequence could be that both hands would leave the steering wheel, a somewhat dubious and unusual maneuver.

This example shows that a parallel system operating just according to Assumption 3 fails to ensure that at least one action system can control all the effector units it requires. Yet unless this condition is realized, no satisfactory action can ensue. Assumption 4 is added to solve this problem, as the input to an effector unit from the dominant action system will now be much stronger than other inputs. It would also avoid the potential cybernetic problem that some action systems might not be physically independent in the brain. The second part of Assumption 4 is derived from ideas on lateral inhibition in sensory systems (e.g., Ratliff, 1965) and those on inhibitory requirements of cell assembly theory (Milner, 1957). It also removes the necessity of having a separate system which selects which action system is to be dominant; the two functions of selection and operation can be performed by the same system. To prove that it is sufficient to produce the required property of always allowing one action system to be dominant clearly depends on a more precise formulation of the action system interactions. A very much simplified formalization of the interaction of action systems, which has the required property, is given in the Appendix.

It is now possible to differentiate between action systems and effector units, solely in terms of their interrelation with other units.

5. An action system has the capacity to inhibit all other action systems sufficiently to become dominant; an effector unit does not.

This would not require a dichotomy of amounts of inhibition that can be produced by systems. With a continuous range of amounts of inhibition, systems only producing a very small amount would never be capable of becoming dominant given normal activity in other systems.

6. An action system can have two excitatory types of input, that necessary for its ongoing operation (specific input) and that necessary to activate it and set its goal (selector input). The selector input can come from either perceptual or motivational systems or other action systems.

For a routine to be entered in programming, the main program must set a number of variables and then use an instruction which passes control to the routine. These variables will determine how the routine operates on that particular entry. For action systems to operate in an analogous way, other systems must activate it and set its goal; these are the functions of the selector input. By contrast, a routine can call for further input of data while operating; this corresponds to the specific input. The difference between these two types of input was noted for skills by Bernstein (1967). For instance when reaching for an object, the visual and kinesthetic input which provides an estimate of the position and velocity of the hand has a different function from that input producing the initial object recognition which determines what action system will be most strongly activated and what its goal will be on that particular operation. The specific input, then, is required for the continual estimation of rapidly changing variables necessary for the operation of the action system; the selector input, for differentiating between this particular operation of the action system and others. The memory requirements for the specific input will normally be small, and insofar as the action system can be modeled by a linear system, it will be zero. The selector input, though, must clearly be stored in a memory for, if the goal and activation of an action system depended on continual input, then an animal could not even turn away from its goal to achieve it.

It would seem possible that only the amount of activating boost and the characteristics of the goal would need to be maintained in the memory; not the input that gives rise to them. However consider the example of braking to an emergency stop in a car. The goal of this action is at one level to have one's foot pressed as firmly down on the brake as possible. At another level, it is to stop the car as quickly as possible. Yet what is done after the car is stopped must depend on why the stop was initiated even if the perceptual input that initiated it is no longer present. Thus it seems necessary for the initial selector input to be stored in memory so that the goal of later action systems can be set. There would be no point in storing the specific input for this purpose, since it is specific to the action system.

7. When an action system becomes dominant, its selector input is stored in a memory, so that it, too, can act as a later selector input.

Finally, for humans, at least some action systems must be capable of being activated and having their goals set by speech input; otherwise speech could not influence action. Thus speech must produce one form of selector input. Similarly in order to have a need satisfied by another person, it must be possible to transmit by means of language what action system is dominant (at least for some action systems), and what is its goal.

8. A speech system exists with both input and output aspects. As input, it can produce another form of selector input. As output, it acts as a nondominant action system symbolizing which action system is dominant and what is its goal.

EMPIRICAL TESTING

It is difficult to test complex information-processing theories using input–output relations. Yet a test of the present theory involves a problem other than complexity. For most information-processing theories, the units are known in advance. For example, testing Morton's (1969) "logogen" theory depends on the reasonable assumption that to each word in the language there corresponds a logogen. Thus, Morton's theory can be tested with its units—the logogens known in advance from other sources. For the present theory, the units—the action systems—cannot be established in this way. No dictionary of action systems exists. Since skill or cognitive competence does not stabilize either over individuals or within an individual's life, unlike language competence, even the prospect of a dictionary seems remote.

The problem of empirical testing is approached in two ways. It is shown that the theory offers an answer to certain known difficulties for existing theories. Second, one situation is examined to show that the theory could be corroborated more closely, if appropriate assumptions about action systems are made.

Two widely held generalizations have been implicitly assumed in developing this theory. The first is that motor skills can be simulated in analogue fashion (e.g., Gibbs, 1970; Wiener, 1948). Bernstein (1967) produced evidence for the equivalent generalization that skilled performance can be fitted by a set of differential equations. The second is that thought can be partially simulated by serially organized computer programs (e.g., Arbib, 1969; Miller et al., 1960; Newall et al., 1958; and in essence, Selz, 1922).

The other widely held generalization about skill is the analogy between the nervous system and a limited-capacity information-transmitting channel. While no longer accepted in its strict mathematical sense, it remains a useful rough description of certain aspects of skilled behavior (Broadbent, 1971; Welford, 1968). On the present theory, this characteristic is deduced by considering how a system based on the first generalization, that is, motor skills being simulated in analogue fashion, could operate in practice, for the second generalization corresponds to Assumption 4—that only one action system can be strongly activated at any one time.

Similarly it is widely held that thought cannot just be a serial process under control of a master program (Neisser, 1963; Reitman, 1965; also Humphrey, 1951, for a criticism of Selz). On the present model, it is possible to deduce this nonserial characteristic from the way in which nondominant action systems can operate at the same time as a dominant one. In its simplest form, this corresponds to the way in which a simple task can be performed in parallel with a complex one. One can both walk and talk.

More interestingly, the theory is able to account for those changes in dominance which would be difficult to explain in terms of a totally serial program. De Groot's (1965) application of Selz's theory of thought to chess is possibly the most detailed examination of the application of a purely serial theory. He stated:

> The more microscopically we examine the phase structure, the more we come across automatic processes and sequences that must be governed by an extensive series of fixed linkages. Calculating a variant and working out a plan are, in fact, operations which in any experienced player are completely controlled by routine [p. 274].

Yet even he is forced to postulate insightful advances:

> In both thought processes [in two different protocols] we are concerned with "calculational serendipity," that is during routine calculations there may be a sudden coincidentally evoked means-abstraction with respect to the main goal [p. 278].

In the present approach, a chance perceptual configuration could so strongly activate another action system so as to make it dominant. Such a possibility would become more likely, if two additional assumptions hold. First, that selector input can be internally generated by inner speech or imagery; second, that if an action system has not attained its goal, it will remain submaximally activated even when not dominant. However the major point is that the present approach has this potential; a purely serial approach does not.

As an illustration of the way in which the theory can be tested more directly, given assumptions about the action systems operating, consider the assumption from which the deductions about the two generalizations mentioned above flow; namely, that no more than one action system may be dominant at any given time. To test this, it is necessary to examine the effect of simultaneously combining two problems, each requiring a dominant action system for its solution, but which do not require the same effector units or lower level action systems.

One such situation appears to be that used by Posner and Keele (1969). Their subjects performed a reaction time task at varying stages of a task involving a motor movement to a target. The reaction time task is slowed by being performed at the same time as the movement task, which is itself relatively unaffected. This fits with the movement task being controlled by a dominant action system and the reaction time task by a nondominant one. As expected in the theory, the slowing is greatest and independent of target size in the initiation of the movement task, when its action system would need to take in selector input and must be dominant. This is not just a question of perceptual filtering, as shown by other experiments of Posner and Boies (1971) on probe reaction time in perceptual matching tasks where response selection but not perceptual encoding is found to produce slowing. To account for the other variations in slowing over the duration of the movement task would require more detailed assumptions about the change in activation over time of the dominant action system, and in particular the interaction of this with the requirements of the task.

PHENOMENOLOGICAL CORRESPONDENCES

Our mental concepts could not arise were it not for the structure of the human brain. "Smell" depends on the existence of an olfactory system; "image" depends on the possibility of internal input to the higher levels of the perceptual systems. It is held that similar considerations apply also for the concept "consciousness." In the present information-processing model, the brain is a set of relatively independent systems, whose operations are coordinated at any moment by the selector input to the dominant action system. The selector input selects which action system is to be dominant, sets the goal of the action system, and is itself preserved in memory. It is suggested that this input corresponds to the concept consciousness, in the mental-state sense of the word. Thus, when one is conscious of X, a certain action system is receiving selector input, which represents X. The selector input has the dual functions of (*a*) providing the activation that enables the action system to become dominant and of (*b*) setting its goal. To mix levels of analysis, these functions can be said to be the dual functions of consciousness. Such a view of consciousness is related to the one held by Armstrong (1968), according to which it is " a mental state of a person apt for bringing about a certain sort of behaviour [p. 82]."

If consciousness is identical with the selector input to the dominant action system in the sense that a section of a DNA molecule is identical with a gene, then one should be able to demonstrate corresponding properties of the two concepts.

Mental states, with the possible exception of emotions like "free-floating" anxiety, have been held by philosophers since Brentano (1874) to have the property of intentionality, to point beyond themselves, to use Armstrong's (1968) expression; a percept, a memory, an intention are all of something. The selector input, too, points beyond itself, since being used to form the goal of an action system, it refers to only a part of the world.

The model can also explain the different types of conscious content. Some are obvious, for example, percepts would correspond to selector inputs from the perceptual system; intentions would be selector inputs from a previously dominant action system. The less obvious correspondences are those related to lack of conscious content. Since an action system only requires selector input at the beginning of its operation, it can operate with little or no selector input. As known to the Würzburg school (see Humphrey, 1951), the whole of the solution period of a problem is not necessarily filled with conscious content. Moreover, it explains why one is not conscious of processes connected with the lower levels of skills, as these would always be directly controlled by never-dominant effector units. It also helps to explain why skills, initially performed under conscious control, can be performed later without conscious control. If one accepts the idea that learning a skill involves the formation of a neural system of increasing redundancy, then initially the corresponding action system would require maximum activation; with practice it would function satisfactorily with less.

The present approach requires that mental terms in phenomenal reports refer to conscious states. Since conscious states are unobservable, it might be argued that it would be impossible for a child to learn such terms. This would only be plausible if the unobservable events did not have a relatively consistent relationship with observables. Skinner (1957) gives a series of examples of how such learning could occur, if there is a consistent relationship; these examples are not critically dependent on his general behavior theory. In the model, relatively consistent relationships exist between conscious states and observables—the relation, determined by the perceptual system, between a stimulus and the selector input and, more importantly, the control relationship between selector input and action. While the detailed processes by which mental terms attain their complex context-dependent meanings are poorly understood, such meanings do not seem to provide difficulties in principle for the present approach.

DISCUSSION

It has been argued in this paper that consciousness is a scientifically acceptable concept, as it refers to a particular type of brain process and can be mapped into an information-processing concept. This also applies to phenomenological terms such as "percept" and "intention," if used in the appropriate context. It has not been the aim of this paper to provide a complete account of consciousness, as the model besides being incomplete, is highly tentative, as the problem of empirical corroboration remains severe. Moreover, the cybernetic arguments depend on the acceptance of computer programming analogies and on the use of loose criteria such as the ease of neural construction. The incompleteness of the model is shown by no correspondence being given for many types of experience such as the consciousness of relation or doubt or expectation. Nor have aspects of consciousness such as the complexity of conscious perception been considered.

More serious, perhaps, is that the model appears to give no account of certain perceptual types of consciousness seemingly unrelated to action, such as noticing a color or seeing someone's face you half recognize. One possible explanation is that an action system exists which maximizes perceptual input from a part of the world determined by its selector input, combining both motor movements (e.g., eye movements) and resetting of the perceptual system. This process would fit with the orienting response and is analogous to Dixon's (1971) suggestion that when a stimulus comes into consciousness, the "gain" of those perceptual systems relevant to it are increased and of irrelevant ones inhibited. Here, a phenomenological fact could be accounted for by extensions to the model which are related to ideas already in the literature. It illustrates the way in which, if the general correspondence theory is correct, properties of phenomenal experience could be used to help develop information-processing ideas.

The major point of the paper, however, is not to argue for this specific model, but to attempt to demonstrate that a model can be developed with an element isomorphic to consciousness. Such an isomorphism makes comprehensible how consciousness, a concept so basic that it is impossible to define using ordinary language, or even to show, can be part of our scientific conceptual framework. The paper, thus, attempts to justify the increasing practice of using phenomenal reports and concepts in investigations primarily orientated toward the development of mechanistic theories. The mixing of these two types of explanation has been challenged by both behaviorists and phenomenologists (e.g., Giorgi, 1970), who would each accept only one type as legitimate. Eventually, if such a model is well supported empirically, it could provide a foundation for developing theoretical criteria for the interpretation of phenomenal reports in cognitive experiments, where at present only intuitive methods can be used.

REFERENCES

Arbib, M. A. Memory limitations of stimulus-response models. *Psychological Review*, 1969, **76**, 507–510.

Armstrong, D. M. *A materialist theory of mind.* London: Routledge, 1968.

Attneave, F. Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 4. New York: McGraw-Hill, 1962.

Beckner, M. *The biological way of thought.* New York: Columbia University Press, 1959.

Bernstein, N. A. *The co-ordination and regulation of movements.* Oxford: Pergamon Press, 1967.

Borst, C. V. *The mind-brain identity theory.* London: Macmillan, 1970.

Brentano, F. *Psychologie vom empirischen Standpunkte.* Berlin: Duncker & Humboldt, 1874.

Broadbent, D. E. *Perception and communication.* London: Pergamon Press, 1958.

Broadbent, D. E. *Decision and stress.* London: Academic Press, 1971.

De Groot, A. D. *Thought and choice in chess.* The Hague: Mouton, 1965.

Deutsch, J., & Deutsch, D. Attention: Some additional theoretical considerations. *Psychological Review*, 1963, **70**, 80–90.

Dixon, N. F. *Subliminal perception: The nature of a controversy.* Maidenhead, England: McGraw-Hill, 1971.

Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationism and the concept of perception. *Psychological Review*, 1956, **63**, 149–159.

Gibbs, C. B. Servo-control systems in organisms and the transfer of skill. In D. Legge (Ed.), *Skills.* London: Penguin, 1970.

Giorgi, A. *Psychology as a human science: A phenomenologically based approach.* New York: Harper & Row, 1970.

Gray, J. A. The mind-brain identity theory as a scientific hypothesis. *Philosophical Quarterly*, 1971, **21**, 247–254.

Griffith, J. S. On the stability of brain-like structures. *Biophysical Journal*, 1963, **3**, 299–308.

Humphrey, G. *Thinking: An introduction to its experimental psychology.* London: Methuen, 1951.

Malcolm, N. *Knowledge and certainty: Essays and lectures.* Englewood Cliffs, N.J.: Prentice Hall, 1963.

Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston, 1960.

Milner, P. The cell assembly: Mark II. *Psychological Review*, 1957, **64**, 242–252.

Morton, J. Interaction of information in word recognition. *Psychological Review*, 1969, **76**, 165–178.

Natsoulas, T. What are perceptual reports about? *Psychological Bulletin*, 1967, **67**, 249–262.

Natsoulas, T. Concerning introspective knowledge. *Psychological Bulletin*, 1970, **73**, 89–111.

Neisser, U. The multiplicity of thought. *British Journal of Psychology*, 1963, **54**, 1–14.

Newell, A., Shaw, J. C., & Simon, H. A. Elements of a theory of human problem-solving. *Psychological Review*, 1958, **65**, 151–166.

Piaget, J. Piaget's theory. In P. H. Mussen (Ed.), *Manual of child psychology.* (3rd ed.) New York: Wiley, 1970.

Posner, M. I., & Boies, S. J. Components of attention. *Psychological Review*, 1971, **78**, 391–408.

Posner, M. I., & Keele, S. W. Attention demands of movements. In *Proceedings of the 16th internatonal congress of applied psychology.* Amsterdam: Swetts & Zeitlinger, 1969.

RATLIFF, F. *Mach bands: Quantitative studies on neural networks in the retina.* San Francisco: Holden-Day, 1965.

REITMAN, W. R. *Cognition and thought.* New York: Wiley, 1965.

ROSNER, B. S. Psychophysics and neurophysiology. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 4. New York: McGraw-Hill, 1962.

SELZ, O. *Zur Psychologie des produktiven Denken und des Irrtums.* Bonn, Germany: Cohen, 1922.

SHAFFER, J. Could mental states be brain processes? *Journal of Philosophy,* 1961, **58**, 813–822.

SKINNER, B. F. *Verbal behavior.* New York: Appleton-Century-Crofts, 1957.

SMART, J. J. C. Materialism. *Journal of Philosophy,* 1963, **60**, 651–662.

SPERLING, G. Successive approximations to a model for short-term memory. *Acta Psychologica,* 1967, **27**, 285–292.

WAUGH, N. C., & NORMAN, D. A. Primary memory. *Psychological Review,* 1965, **72**, 89–104.

WELFORD, A. T. *Fundamentals of skill.* London: Methuen, 1968.

WIENER, N. *Cybernetics, or control and communication in the animal and the machine.* New York: Wiley, 1948.

WITTGENSTEIN, L. *Philosophical investigations.* Oxford, England: Blackwell, 1953.

WOODSWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology.* New York: Holt, 1954.

APPENDIX

What is required is to show that mutual inhibition between action systems is sufficient to produce the desired dominance property. Since action systems are neural systems, the obvious mathematics to use is neural net theory. However neural net theory is notoriously intractable if applied to complex networks (see Griffith, 1963).

A relatively simple mathematical formalization of the system can be obtained which is tractable and, for one isolated action system, approximates to Griffith's (1963) results for the behavior of a single purely excitatory net. This is to describe the behavior of a system of n action systems by the following set of equations:

$$\frac{dy_i}{dt} = \begin{cases} ax_i + by_i - c(\sum_{j \neq i} z_j) - e & \\ \qquad\qquad \text{if } y_i > 0 \text{ or } \dfrac{dy_i}{dt} > 0 & , \\ 0 \qquad\qquad \text{if } y_i = 0 \text{ and } \dfrac{dy_i}{dt} \not> 0 & \end{cases}$$

where $z_i = \max\ (y_i - d;\ 0)$ and a, b, c, d, e are constant and positive.

In these equations, y_i (nonnegative) represents the activation level of the ith action system, x_i represents its selector input, and z_i represents its contribution to the inhibition of other systems. If one assumes that $c > b$ (i.e., inhibition is the more powerful at high levels of activation), and that if after a certain point in time the selector inputs are constant and at least one is greater than some critical level, then it follows that one and only one of the action systems will have an activation that will increase rapidly and continually. In fact this is a formalization of the mutually inhibitory interaction of a system of pure excitatory neural nets, but the dominance relation should be robust enough to survive complicating the operation of the individual interacting elements.

(Received July 13, 1971; revision received May 9, 1972)

Psychological Review
1972, Vol. 79, No. 5, 394–409

REINFORCEMENT, EXPECTANCY, AND LEARNING[1]

ROBERT C. BOLLES[2]

University of Washington

A variety of recent evidence indicates that contingent reinforcement is neither a necessary nor a sufficient condition for operant learning. This dilemma reopens the old question of What is learned? It is proposed that what laboratory subjects characteristically learn is not a response to a stimulus but rather two kinds of expectancies. One kind of expectancy (S–S*) corresponds rather accurately with environmental stimulus-outcome contingencies, while the second (R–S*) is a less faithful representation of response-outcome contingencies. The reinforcement procedure merely permits both kinds to be learned.

One of psychology's most recent contributions to the general welfare has been the principle of reinforcement. Increasing numbers of people have become familiar with the basic idea of reinforcement and have begun to apply this idea to the solution of important practical problems. Increasing numbers of people have even begun to see the principle of reinforcement as providing a new way to explain behavior. It is not the purpose of this paper to question this contribution. The value of the reinforcement *procedure* will be judged in time by its practical validity, that is, by whether it permits the psychologist, or anyone else who cares to use it, better control over the behavior of organisms. The concern here is with the assumption so frequently made by psychologists that when the reinforcement procedure produces a change in behavior, such a change must be caused by an underlying reinforcement *process*. This hypothetical reinforcement process is to be judged not in practical terms but in terms of the generality, internal consistency, parsimony, and empirically confirmability of the resulting laws of learning. Considering these criteria in the light of recent developments indicates that the concept of a reinforcement mechanism cannot be taken for granted. The possibility of learning without an underlying reinforcement mechanism may appear unthinkable to some, but this is what the present paper proposes.

We begin by surveying some of the difficulties currently confronting the reinforcement concept. Then we consider some alternatives to reinforcement as the fundamental basis of learning. Two specific alternatives will be considered: an incentive motivation approach and a cognitive approach. Some of the potential advantages and difficulties with these alternative approaches are described. But first, let us justify this undertaking by noting some recently discovered failures of the hypothetical reinforcement process to account for what happens in certain learning situations.

SOME FAILURES OF REINFORCEMENT TO CONTROL BEHAVIOR

The principal evidence supporting the assumption of a reinforcement process is that the reinforcement procedure often does work. If certain events are made contingent upon the occurrence of a particular response, and consequently that response becomes stronger, then the change in behavior can be theoretically attributed to some kind of response-strengthening mechanism or process. But what assumptions can be made when an event which has been found to be a reinforcer under other circumstances is regularly applied to a response, when the required response occurs at a reasonably high level, but when the reinforcement contingency fails to produce an increase in response strength? Such cases, and there is a growing number of them, indicate that the reinforcement procedure is not sufficient to produce learning and that the assumed reinforcement process can fail to operate when it should. There also are cases which present the converse difficulty. That is, there are situations in which an animal acquires a new response too readily—before the reinforcement procedure can be effectively imposed. Such cases indicate that the reinforcement procedure is not necessary for learning and that learning can occur in the absence of the assumed reinforcement process. These are very illuminating cases. Let us consider some of them in a little more detail.

---

[1] Supported by National Science Foundation Research Grant GB-20801.

[2] Requests for reprints should be sent to Robert C. Bolles, Department of Psychology, University of Washington, Seattle, Washington 98195.

Superstition

Superstitious behavior is frequently observed when an animal is provided with food on a schedule, such as once a minute, independently of its behavior. In the absence of a reinforcement contingency, there is nothing the animal has to do; it could simply relax and wait for the inevitable arrival of food. But typically the animal acquires stereotyped, repetitive segments of behavior, which may ultimately come to fill most of the time interval between presentations of food. The reinforcement theorist can argue that any behavior which tends to occur immediately prior to reinforcement will be reinforced, and that once such behavior starts, a long homogeneous chain of it will be built up backwards from the time of reinforcement. Therefore, even though reinforcement is not contingent upon this behavior, the reinforcement theorist can still assume reinforcement to be effective in establishing it. Behavior that appears at first to be anomalous can be treated as yet one more sign of the effectiveness of reinforcement in producing learning. This reinforcement account of superstitious behavior has stood virtually unquestioned since Skinner first proposed it in 1948.

Staddon and Simmelhag's (1971) recent analysis of superstitious behavior indicates, however, that this traditional account of the phenomenon is inadequate on many counts. They report that there are two classes of behavior occurring during the interfeeding interval. One class occurs just prior to reinforcement; it tends to be directed toward the food hopper, and it resembles in many respects the behavior which accompanies feeding. Moreover, reinforcement appears to be effective in producing this behavior near the end of the interfeeding interval. But there is also a class of what Staddon and Simmelhag call "interim" behavior which appears to follow other laws. It is topographically different, functionally different, and also different in the sense that it does not appear to be related to the upcoming reinforcement. The occurrence of the two classes of behavior appears to be governed primarily by the temporal relationship with reinforcement, that is, by whether or not the time intervals in which they occur are predictive of food. Thus, superstitiously reinforced behaviors occur in the interval immediately preceding food, and interim behaviors occur during time intervals correlated with the absence of food. Therefore, interim behavior is isolated from any reinforcement mechanism, first by lack of contiguity, and more importantly, by the intervention of other behavior. Thus, it cannot be merely part of a chain of responses which is superstitiously reinforced.

Polydipsia

A similar case can be made against a reinforcement interpretation of the excess drinking which occurs when rats are given food on fixed-interval schedules. This polydipsia has been extensively studied in the past few years and we now know of a number of parameters that encourage or discourage its occurrence (Falk, 1969). Although there is still no completely adequate hypothesis to explain this excessive drinking, it is clear that the behavior is not just maintained superstitiously. One reason for this conclusion is that just as Staddon and Simmelhag found for the interim behavior occurring in the situation they studied, the drinking never immediately precedes reinforcement. Therefore, there is no way in which the behavior could have become chained to bar pressing or other previously reinforced behavior. Actually the drinking appears to be elicited by the previous eating. Here we have a second example of behavior which appears superficially to be due to reinforcement but which further analysis shows is not. Rather, it appears to occur because the animal is hungry and expects food but the immediate cues in the situation predict the momentary absence of food.

Misbehavior

The reinforcement procedure can be used so successfully to establish long chains of unusual responses that we have come to take such applications for granted. Sometimes, however, these procedures fail to work. Breland and Breland (1961) have described some dramatic failures of animals to learn responses upon which reinforcement is contingent. In one case, a raccoon was supposed to carry a couple of coins over to a piggy bank and deposit them. Success seemed assured because the raccoon was intelligent and cooperative; it had great manual dexterity, and it was easy to train it to carry a single wooden coin over to the piggy bank. But when the response requirement was shifted so the raccoon had to deposit two coins, the learned behavior began to break down. With continued training the situation worsened. The raccoon had its own behavior which intruded; it would rub the coins together and dip them in and out of the piggy bank. This intruding behavior became quite persistent, even though it greatly delayed the reinforcement that might have explained its occurrence. The raccoon's intruding behavior came to predominate to such an extent that training had to be abandoned. Breland and Breland discuss other failures of reinforcement in which intruding behavior is so prevalent that the animal cuts itself off entirely from the original reinforcement contingency. Breland and Breland did not attempt to explain the phenomenon, except to note that with continued training, certain topographic features of the animal's natural food-getting behavior come to replace the arbitrary food-getting behavior on which the reinforcement is experimentally contingent. So although a reinforcement contingency can be used to establish and maintain some kinds of behavior, in the cases that Breland and Breland wrote about, it could not successfully prevent a drift toward instinctive forms of food-getting behavior that are characteristic of a given animal. The more the reinforcement contingency was imposed, the less successful it appeared to be in controlling the behavior.

Auto Shaping

The hypothesis that reinforcement is a necessary part of the learning process falls into difficulty not just in these cases of poor learning. It has difficulty in other cases where learning is too good. One example is the recently discovered phenomenon of auto shaping (Brown & Jenkins, 1968). Part of the lore of the beginning student's animal laboratory is that the desired response is initially shaped up in the operant situation through the careful and gradual application of reinforcement contingencies. Initially, the animal is magazine trained, that is, it is trained to take food at the sound of the feeder mechanism or the lighting of the food hopper. Then this secondary reinforcement together with the accompanying primary reinforcement (food) is used to shape up increasingly good approximations to the desired bar-press or key-peck response. This gradual shaping process ordinarily takes the student about an hour, by which time the animal is usually under the direct control of the automated programming apparatus and is busily making the required operant response. The recent study by Brown and Jenkins (1968) showed, however, that a lot of psychology students have wasted a lot of their time, because none of the magazine training, none of the gradual response shaping, and none of the experimenter intervention into the automatic programming is necessary. What Brown and Jenkins did was to put the key-peck reinforcement contingency into effect at the outset and from time to time give animals free food (food not contingent upon the animals' behavior). After an hour on this program, a bird is under the direct control of the automated programming apparatus and busily pecking the key to obtain reinforcements while the experimenter has been free for the last hour to seek his own.

Brown and Jenkins' finding by itself does not pose a serious challenge to reinforcement theory because it could be argued that the key-peck response has a very high operant rate in the hungry pigeon. Since any accidental key pecks are reinforced, the basic contingency has a good chance of operating upon the behavior and establishing it rather quickly. Successful auto shaping could presumably be further insured by making the key interesting, for example, by illuminating it, and by placing it close to the food hopper, etc. But a quite different interpretation is demanded by Williams and Williams' (1969) clever extension of the auto-shaping experiment. In their study, free food was delivered from time to time independently of the bird's behavior but the key-peck reinforcement contingency was made ineffective. Actually, whenever the bird pecked the key the next scheduled presentation of food was delayed, so Williams and Williams' birds could have received food more frequently by not responding than by responding. But in spite of this negative correlation with reinforcement, the pecking response became so firmly established and occurred at such a high rate that some animals cut themselves off from all subsequent reinforcement (see also Neuringer, 1970). The results are, therefore, very much like those reported by Breland and Breland in which an animal's innate food-getting behavior, pecking in this case, is acquired *in spite of* the reinforcement contingency. The reinforcement contingency is evidently irrelevant in such cases.

Still further complications arise from some results recently found by Moore (1971). Moore made detailed high-speed analyses of the topography of the key-pecking response in pigeons auto shaped with food and pigeons auto shaped with water. The topography of the consummatory responses is known to be different in the two cases: eating is an isolated, sharp, hard peck with the beak partially opened as if seizing a morsel of food; the drinking response (the pigeon is unique in taking water the same way mammals do, rather than in the typical avian manner) consists of a series of brief gentle nibbling or mumbling reactions. Moore found that when the key-peck response became established in the auto-shaping situation, it bore a striking resemblance to the appropriate consummatory response. He concluded that in this situation the key-peck response looks more like a classically conditioned consummatory response than an operant. We suddenly have the specter of what must be the single most-studied operant response not being an operant and presumably not being acquired by operant principles!

Each of the cases that have just been considered illustrate an animal learning—or not learning—some response in spite of the basic law of operant learning, that is, the empirical law of reinforcement. But even though the animal's behavior does not abide by the law of reinforcement in these cases, it is nonetheless lawful. The behavior is not capricious; it actually demonstrates a remarkable consistency. In each case the observed behavior is specific to the given species, and in each case it is specific to the given motivational state. Thus, the hungry rat looks and behaves like a hungry rat whether or not it's behavior is governed by response-contingent food. Breland and Breland's hungry raccoon rubbed and dipped its food tokens just like a hungry raccoon. Staddon and Simmelhag's hungry pigeon behaved like a hungry pigeon—it scratched, pecked, cocked its head this way and that, and pecked at interesting things in the world around it. Moore's pigeons behaved as though they were literally eating or drinking the response key. Thus, even though these different behaviors failed to be governed by the law of reinforcement, they appeared to follow a higher law: in each case the animal acted as an animal of its species acts when it expects food.

The same general law apparently also applies in those cases where a reinforcement contingency is effective in controlling the animal's behavior; that is, in the great bulk of situations that have been studied in the laboratory. For example, if we train a rat to run down an alley for food, we are likely to find after a few trials that it gives many signs of eager expectancy. It greets us eagerly when we initiate a trial; it eagerly pushes at the start-box door; it eagerly runs down the alley overcoming all obstacles and being frustrated by all delays along the way. There is nothing extraordinary or magical about such behavior. Indeed, it seems to be characteristic of all learning situations. Guthrie (1935) and Hull (1931) attributed such behavior to the same learning processes which they hypothesized to explain the running response itself.

Eager behavior, expectant, or anticipatory behavior has hardly ever been studied in its own right. It has been typically regarded as little more than a by-product of the primary learning process. This second-class status would seem to be rather inappropriate in view of the ubiquity of expectancy effects. There is evidence that the animal expects food in those cases where reinforcement fails to work as well as those cases where it does. One purpose of the present paper is to elevate this idea—that animals can anticipate important events such as food and electric shock—to a first principle of learning. Before proceeding with this development, however, let us look at another area of animal learning where the application of reinforcement principles has been something less than successful.

AVOIDANCE LEARNING

Avoidance learning has proved to be a perennial challenge to learning theorists. The chief problem it poses at the present time is that a reinforcement interpretation of avoidance learning appears to be much more believable for some avoidance responses than it is for others. For example, a variety of studies have attempted to assess the importance of CS termination in avoidance learning, and it has been found that while CS termination is relatively effective for some avoidance responses, such as running in a shuttle box, it makes little contribution to the acquisition of other responses, such as running in a wheel (e.g., Bolles, Stokes, & Younger, 1966). In the case of a response like bar pressing, neither contingent CS termination nor anything else seems to provide very effective reinforcement (this literature is reviewed in more detail in Bolles, 1972). A further outcome of these studies with different avoidance responses is that the traditional reinforcement contingency, that is, CS termination, appears to have the least effect in those situations where the avoidance response is most rapidly acquired. Thus, in one-way avoidance, which is very rapidly learned by the rat, there is little effect of CS termination or even the use of the CS as a warning stimulus. The rat performs quite well without any CS. Thus it does not seem reasonable to claim, as a reinforcement theorist might, that CS termination is ordinarily the source of reinforcement for avoidance learning, but for some reason it does not work in some situations.[3]

Essentially, the same situation has been found for the contribution of shock escape to avoidance learning (Bolles, 1972). It contributes importantly to the acquisition of some avoidance responses and has little or no effect upon the acquisition of others. Thus, the fundamental problem for the reinforcement theorist is that when the various experimental procedures that are ordinarily made a part of the avoidance learning situation are experimentally isolated, none of them appears to have a consistent effect. In predicting the outcome of an avoidance learning experiment, it is necessary to know not only what the experimental contingencies are, but also what response is required of the animal. As an empirical generalization, this conclusion leaves us in the peculiar and unsettling position of having some responses learnable by a given animal and others not, and it leaves the reinforcement theorist in the uncomfortable position of having his allegedly universal laws of reinforcement effective in some cases and not effective in others.

One way out of the puzzle is to argue (Bolles, 1970) that avoidance learning does not require a reinforcement process, but is based on an animal's innate defensive behaviors. All that is required is to restrict the animal's behavior to its species-specific defensive reactions and then institute an avoidance contingency to punish inappropriate defensive reactions. To illustrate the argument, consider the rat and suppose that the rat's only defensive reactions are freezing and running away. Then if the experimental situation is arranged so that running away is an acceptable avoidance response, all that is needed to produce this behavior is to punish freezing. Such punishment is conveniently effected by the occurrence of unavoided shock. Under these conditions it is not necessary to train or reinforce running away. The learning which occurs under these conditions is parsimoniously explained without any reference to reinforcement. The response is already there, and will occur if the other response, freezing, is punished. How rapidly an avoidance response is learned becomes, in this view, a function of how compatible the response is with the animal's species-specific defense reactions. Thus, if an animal is required to learn a response which is not in its repertory of natural defense reactions, then we can expect slow learning and poor performance. On the other hand, if the animal is required to learn a response which is part of its natural defense repertory, then that behavior should be promptly acquired.

---

[3] Another failure of reinforcement in avoidance behavior is found in extinction. While CS termination can help maintain established avoidance behavior (Katzev, 1967), this appears to be true only under certain special conditions, namely, total absence of further shock. Whether or not the animal has the advantage of the presumed reinforcing effects of CS termination accounts for very little of the total variance over the range of different possible extinction conditions (Bolles, Moot, & Grossen, 1971).

Notice the parallel between this interpretation of avoidance behavior and the conclusion previously drawn about the failure of reinforcement in appetitive situations. In each case, the animal comes into the experimental situation with certain species-specific behaviors; food-getting responses on the one hand and defensive reactions on the other. In both cases, the chief effect of putting the animal in the experimental situation is to produce an expectancy; an expectancy of food in the one instance and an expectancy of shock in the other. The animal then gives us its characteristic behavior; food-getting responses or defensive reactions. Before proceeding with the systematization of these ideas, let us look at some alternative schemes for minimizing or eliminating the hypothetical role of reinforcement in animal learning.

INCENTIVE MOTIVATION MODELS

The reinforcement procedure is commonly assumed to have two quite different kinds of behavioral effects. One is the direct associative strengthening of the contingent response through reinforcement; the second is an indirect, motivational effect produced through some kind of "second process." If an animal is given food for running down an alley and into a goal box, the presentation of food is said to reinforce running and also to attach incentive motivation to whatever stimuli are near to or similar to the goal box stimuli, so that the animal will run faster whenever these stimuli are subsequently presented. Moreover, if a lot of incentive motivation is attached to one cue by giving reinforcement in its presence, and if none is attached to a second cue by withholding reinforcement in its presence, we may have an animal choosing or approaching the first cue, or responding more vigorously to it, even though there is no difference in the associative strengths of the responses evoked by the two stimuli.

Thus, the idea that incentive motivation can be differentially associated with different cues accomplishes much of the same conceptual work in the learning experiment that reinforcement is supposed to accomplish. A natural question then is the extent to which a change in response strength generally reflects the effects of associative strengthening and how much of it is attributable to changes in incentive motivation. Mowrer (1960) was one of the first to take the fateful step of attributing all of the response-strengthening effects of the reinforcement procedure to incentive motivation, leaving none of it to be explained by associative strengthening per se. According to Mowrer, one reason the incentive mechanism is such an effective mimic of the reinforcement mechanism is that the proprioceptive feedback from the response that leads to reinforcement is itself paired with reinforcement. Then whenever that response begins to occur, its own feedback will provide incentive motivation for its continuance. If the incentive motivation is positive, there will be a positive feedback loop which causes the incipient response to appear in full strength. If, on the other hand, the incentive motivation is negative, then the feedback loop is negative and the incipient response is inhibited.

Following Mowrer's example, other writers have noted the peculiar redundancy in the variety of concepts that are used to explain learned behavior. As Walker (1969) puts it: an instrumental response is first of all *pushed* by whatever drive conditions may exist; it is also *pulled* by the animal's anticipation of reinforcement (incentive motivation); it is also determined by the *structure* of the learning task (the S–R association itself); finally, the response is *glued* to the stimulus by virtue of reinforcement. Walker observes that we have no need for all four of these explanatory concepts, and the one he proposes to give up is reinforcement. Bindra (1969) has made a similar proposal; he suggests that incentive cues arouse a central motive state which triggers or releases fixed behavior patterns directed at the cues. The reinforcement procedure is assumed to have no direct effect on responses, it only establishes incentive cues. Logan (1969) has also suggested that learned behavior is controlled by motivational factors, and that no traditional concept of reinforcement is necessary. Logan only briefly suggests how an incentive motivation factor can initiate behavior; he assumes that stimuli which precede reinforcement tend to elicit "anticipations" of the reinforced response. The common message of these recent writers appears to be that the principles of incentive motivation provide sufficient explanatory power that we do not need any principles of reinforcement. Incentive motivation can explain anything that reinforcement can—and more besides.

The previously described failures of reinforcement to control behavior all presented a characteristic motivational appearance. For example, we may think of Breland and Breland's misbehaving raccoon as having so much incentive motivation associated with the food token that it directed its species-specific food-getting behavior at it. Similarly, we may think of the CS in the avoidance learning situation as producing so much fear that the rat can emit no behavior other than freezing or running away. In each case, the development of these nonreinforced behaviors can be taken as an index of the animal's incentive motivation; positive on the one hand, negative on the other. Because of the ubiquity of the motivational phenomena, an incentive motivational interpretation of the reinforcement procedure promises to have greater generality than the original associative strengthening interpretation of reinforcement. That is, it appears to be applicable to a broader range of phenomena.

Incentive motivation models have other virtues to recommend them. But existing models also present some problems. One problem has been frequently noted as a criticism of Mowrer's (1960) formulation. It is asserted that Mowrer's analysis leaves open the question of how behavior is initiated. Thus if the feedback from the incipient beginnings of a response serves to motivate its occurrence, then we can account for the probability and the vigor of the response, but how are we to explain its beginning? It would appear that situational stimuli must have some minimal amount of associative control over the response, just enough to get it started. But if this is the case, then the concepts of habit and reinforcement which Mowrer so ceremoniously threw out the front door would appear to have slipped in the back door. Bindra (1969) has circumvented this problem by postulating that in addition to providing motivation for ongoing behavior, motivating cues also elicit fixed action patterns of approach or orientation. The real question here is whether incentive motivation simply operates upon existing response tendencies or can it be credited with associative powers, that is, the capacity to initiate responses. The classical treatment of incentive motivation (e.g., Hull, 1952; Spence, 1956) gave it a purely motivational role such that it simply multiplied the habit strengths of ongoing behaviors. But this assumption clearly can be justified only if there are other associative determinates of behavior; that is, some mechanism, such as the reinforcement of S–R associations, to give behavior its directionality. Otherwise it is necessary to follow Bindra and give incentive motivation specific response-eliciting powers.

Recent writers have been reluctant to speculate about the mechanisms of incentive motivation, and one must refer again to the classic models of Hull and of Spence for any tangible proposal of how incentive motivation works. The classical model of incentive motivation is based on three theoretical assumptions. One that was just mentioned is that incentive motivation is a motivator, that is, it multiplies existing habits. Although incentive (K) is assumed to be directive in the sense that it is elicited by particular stimuli and predisposes the animal to make particular consummatory responses, its effect on instrumental behavior is nondirectional. K is just as general as D in that it is assumed not to have selective effects on different behaviors. Thus the classical concept of incentive motivation as developed by Hull and by Spence cannot replace a direct response-strengthening concept.

The second assumption of classical incentive motivation theory is that it is reductively based on some emotional response of the autonomic nervous system. It is assumed to be mediated by the feedback from an appetitive r_g or an aversive emotional r_e. The virtue claimed for this assumption is that because the mediating linkages in incentive motivation would all involve stimuli (Ss) and responses (Rs) (as opposed to expectancies or other less tangible mediators), it is possible to apply the established corpus of knowledge about the learning of Ss and Rs so as to provide some lawfulness for the acquisition of the incentive mediator. In short, although the mediating response is different from learned instrumental behavior, many of the same experimental manipulations can be applied to its general operation because it has the same fundamental S–R character. However, recent reviewers, particularly Rescorla and Solomon (1967), have shown that neither the hypothetical appetitive mediator (r_g) nor the aversive mediator (r_e) can be put into a one-to-one relationship with any observable peripheral response. The attempts to measure salivation or some other autonomic concomitants of hunger have produced, at best, equivocal results. The attempts to correlate heart rate or other autonomic concomitants of fear have been thoroughly disheartening. It can only be concluded that as parsimonious and attractive as the idea of S–R mediation of incentive motivation may be, this assumption has not been, and probably cannot be, empirically justified.

The third assumption upon which classical incentive motivation theory rests is that whether we can observe it or not the mediating response, r_g or r_e, is acquired by a process of classical conditioning. Although a few writers (e.g., Spence, 1956) have considered other possibilities, most theorists seem to have accepted this assumption. The difficulty here is conceptual rather than empirical. The assumption is based on the same kind of failure to distinguish a learning procedure and a learning process that was discussed at the beginning of this paper. It can be granted that the reinforcement procedure works, at least some of the time, but does it make sense to say that operation of the reinforcement procedure can be explained by means of a Pavlovian process when the only evidence for a Pavlovian process is that the Pavlovian procedure works, at least some of the time? To put it another way, the only evidence for a Pavlovian learning process is that the Pavlovian procedure does sometimes affect behavior, but there is no more a priori reason to believe that this is brought about by a Pavlovian S–R strengthening process than to believe the original idea that instrumental learning is brought about by a Thorndikian S–R strengthening process.

Thus, the three assumptions upon which classical incentive motivation theory is based can all be challenged. And whereas no one of these difficulties taken by itself is likely to prove fatal to the theory, it should be clear that the future success of this kind of explanatory model will require a notable departure from the classical incentive motivation model originally formulated by Hull. It should also be clear that a classical incentive model cannot replace a reinforcement interpretation of instrumental learning; it can only supplement it. One alternative model has been suggested by Rescorla and Solomon (1967). It can be best described in terms of an illustrative experiment reported by Rescorla and LoLordo (1965). Dogs were first trained to stable avoidance performance; then they were trained with a Pavlovian procedure in which a CS was inevitably followed by shock; then they were returned to the previous avoidance situation and the Pavlovian CS was introduced briefly to determine its effect upon the previously established avoidance response. The CS had a marked facilitory (motivating) effect. Rescorla and Solomon's (1967) review of the literature indicates that we cannot accept the old idea that a fear response had become classically conditioned to the CS as a result of the Pavlovian procedure. How then are we to explain the motivating effect of the CS? The answer proposed by Rescorla and Solomon, and subsequently endorsed by other writers (e.g., Bolles & Moot, 1972; Trapold & Overmier, 1972), is that the animals did not learn a mediating response in the Pavlovian situation; they learned *the predictive relationship between the CS and shock*. Then when the dogs were returned to the avoidance situation, the CS again produced an expectation of shock and they tended to respond as they would to such an expectancy, that is, by making the learned avoidance response.

The key point in Rescorla and Solomon's analysis is that the Pavlovian procedure—or any other procedure in which classical conditioning is hypothesized to occur—provides the subject with an opportunity to learn about the predictive relationships between environmental events. In effect, animals cannot only learn responses to stimuli, they can also learn expectancies. Perhaps it is these S–S expectancies, rather than mediating responses, that provide the basis of incentive motivation.

This analysis has brought us at last to the old question of what is learned in a learning experiment. We have rediscovered the fundamental assumption which underlies both reinforcement theory and classical incentive motivation theory, namely, that when an animal learns, it must necessarily learn a response. Let us turn now to a different kind of model, a cognitive type of model which is not constrained by this fundamental assumption. Let us see where we can go with incentive theory if we allow expectancies to be learned.

EXPECTANCY MODELS

The idea that learning consists of the acquisition of expectancies is certainly not new. It was the main principle underlying Edward Tolman's life work. Tolman (1949) maintained that animals learn a number of different things in different kinds of situations. But the acquisition of S–R connections which cause the animal to respond in a particular way in a particular situation was always denied to be an important part of what is learned. The basic thing acquired in a learning experiment was always an expectancy. Expectancy was a primitive, undefined term in his system; it was a hypothetical behavioristic construct; it was not mentalistic and it was not teleological; it was behavioristic. Tolman's "expectancy" differed from other behavioristic constructs mainly in the very important respect that it did not require the prior assumption of an S–R unit.

In the remainder of this paper I develop the idea of expectancy in the light of recent experimental work, and I propose a model based on the assumption that what is learned in most learning situations is an expectancy.

Primary Law of Learning

A learning situation is one that imposes upon an animal new environmental contingencies. Learning consists of adaption to these contingencies so that the animal can act appropriately to impending events. Since this would seem to be the most basic aspect of learning, I will maintain that when learning occurs, what is learned is a direct appreciation of the new environmental contingencies. Let me propose the following explicit primary law of learning: What is learned is that certain events, cues, (S), predict certain other, biologically important events, consequences, (S*). An animal may incidently show new responses, but what it learns is an expectancy that represents and corresponds to the S–S* contingency.

The primary law of learning asserts little more than that animals, at least some animals some of the time, can appreciate new environmental contingencies. Information about a new order of things in the environment is processed and stored for subsequent use by the animal. Expectancy is simply a name for this stored information. The linguistic rules for using "expectancy" are essentially those in the everyday language.

Thus, it seems proper to refer to the animal in a particular situation as "expecting" a particular outcome. It also seems proper to speak of the animal as having a particular "expectancy" in a particular situation. The expectancy itself is of course central; it is not directly observable. For the moment, I do not wish to conjecture either about its localization in the nervous system or to what extent it may be represented in consciousness. The more immediate task is to postulate rules so that expectancies can be used to explain behavior.

Secondary Law of Learning

The orderliness of the experimental world imposed upon an animal may include not only predictable sequences of events (stimulus contingencies) but also a contingency between the animal's behavior and programmed consequences. That is, there may be a predictive relationship between the animal's behavior, R, and the consequences of responding, S*. I propose as a secondary law of learning that animals can learn an expectancy which corresponds to this kind of predictive relationship; formally: animals can learn R–S* expectancies that represent and correspond to the R–S* contingencies in their environment. These expectancies, together with S–S* expectancies, constitute all of what is learned in most instrumental learning experiments.

These two kinds of expectancies are verbally somewhat different from the expectancies that Tolman (1932) described. Some of the differences are purely verbal, but others appear to be real. For example, Tolman spoke about the animal acquiring "knowledge" about its environment, whereas I have referred to the representation of information. This difference seems to be purely verbal. On the other hand, the environments in which Tolman ran his animals were typically organized spatially. His experimental situations consisted of tortuous pathways, blind alleys, detours, etc. Accordingly, when his animals learned, they were assimilating spatially organized information and their expectancies embodied this type of information. Tolman himself thought of the learning as knowing how to get to a goal. By contrast, the experiments I ran typically involved events that are organized in time. One cue is followed by shock, another is followed by food. Accordingly, when my animals learn they are acquiring temporally organized information, and the expectancies described here embody this kind of information. In effect, Tolman's animals were always "getting to" some expected outcome whereas mine were typically "waiting-for" some outcome. This difference is a real one, no doubt, but it reflects primarily a matter of emphasis. The rat is probably able to learn and act in accordance with both of these kinds of expectancies—and probably others as well.

The expectancies which Tolman proposed typically included a response term; they were given an S–R–S form. The animal expects a certain outcome as a consequence of responding in a certain way in a certain situation. This kind of element was a handy device to describe representations of the complex environmental contingencies in Tolman's experiments. But an S–R–S formulation clearly does not lend itself readily to the description of what animals learn in experiments in which no particular response is required, for example, in a Pavlovian experiment. Additionally, Tolman (1949) recognized that learning could involve contents other than S–R–S expectancies, but he, as well as subsequent writers, has tended to attach different labels to them. Thus, contents of the R–S* form are often called beliefs; Irwin (1971) calls them act–outcome expectancies.[4] These differences seem to reflect mainly verbal differences.

There is a real question, however, of how many kinds of expectancy should be recognized. I have proposed two kinds to correspond with the two kinds of contingencies that are built into experimental situations. This strategy should provide a more flexible account of what is learned than the one S–R–S* kind Tolman appeared to rely on in his early writing (e.g., Tolman, 1932), without introducing the variety of kinds of learning he distinguished in his 1949 paper. A further reason for having two kinds of expectancies is that, as described below, they seem to be subject to somewhat different principles.

---

[4] Estes (1969) calls them linkages. He suggests that learning can involve three different kinds of associative linkages: the traditional kind between cues and responses, linkages between cues and outcomes, and linkages between responses and outcomes.

Law of Performance

Any theoretical account of behavior must embody some rule for translating what is learned into observable behavior. The following is sufficient: Two expectancies of the form S–S* and R–S* are "synthesized," or combined in a "psychological syllogism" so that in the presence of the cue S the animal is likely to make the response R.

Notice that this is not a statement about learning. Learning involves only the acquisition of the two expectancies that become synthesized. It is simply a statement about how expectancies produce performance. And it asserts merely that they do.

There is a vital point here which needs some elaboration. The point was raised perhaps most dramatically by Guthrie when he made his famous remark that Tolman's rat is so immersed in thought at the choice point that it is quite unable to move a muscle. Guthrie, acting here as a spokesman for all S–R theorists, seems unable to comprehend how having an expectancy can produce movement. An expectancy is merely a hypothetical construct; it is postulated to be an unobservable central event. How can it generate behavior? Guthrie proceeds to tell us that only an S–R association can generate behavior. But what is an S–R association? Can it be anything other than a hypothetical construct, a postulated unobservable, a central event? It is a mistake to think that the "S" and "R" in the S–R association are anything other than theoretical fiction with merely the same name as the S and the R that constitute observable data. We can all observe the animal's behavior in the situation, that is, we can all see the correlation between the situation and the behavior, and we can all see this correlation increase during the course of a learning experiment. But the S–R association has no more prior claim to explaining this correlation than an expectancy or any other theoretical mechanism we might care to construct. The only reason the S–R association explains movement is that it is postulated to do so. An expectancy explains movement because it is postulated to do so.

The great promise and enduring appeal of S–R associationism lies not in the simplicity of its semantics, for this simplicity is an illusion, but rather in the simplicity of its syntax.[5] Guthrie's own analysis required the ad hoc postulation of numerous molecular elements, internal stimuli, and muscle movements, but these elements were assumed to combine in a simple S–R bond that theoretically mimicked the growing correlation between the observable S and R. In practice, his analysis required a vastly complex semantics to preserve the simplicity of the S–R association; Ss and Rs had to be invoked after the fact. Hull's associative unit, habit, preserved the simple S–R association, but embedded it in a more complex syntactical system (Hull, 1952). Even so, the preservation of the theoretical system introduced a mass of semantic problems (Koch, 1954). Cognitive psychology is characterized by its insistence upon a basic syntactical element or process with greater complexity than the S–R bond. The basic syntactical process has been likened to reading a map (Tolman), or, in more recent times, to coding, storing, and retrieving information, or making a decision (Irwin, 1971). The important question is whether this approach will yield a more practical set of semantic rules for relating theoretical constructs to observable data.

Law of Prior Expectancies

The law of performance accounts for the effectiveness of a reinforcement contingency in controlling behavior. But it remains to account for the kind of case described earlier in which the reinforcement procedure fails to control behavior. These cases illustrate the important law of prior expectancies: An animal can be expected to have previously learned or innate R–S* expectancies. In accordance with the law of performance, when the animal has learned a strong S–S* expectancy, its behavior may be dominated by such prepotent R–S* expectancies rather than those which correspond to environmental R–S* contingencies.

---

[5] The syntax of a theory are the rules for relating its theoretical terms to each other. The semantics of a theory are the rules for relating its theoretical terms to observational data.

Prior expectancies impose constraints upon what an animal can learn. Sometimes an animal may be constrained by a prior S–S* expectancy. The innate fears described by Hebb (1946) are a case in point. Stimuli that release fixed-action patterns can be thought of in the same way. But the most common as well as the most constraining expectancies are of the R–S* variety. For example, the hungry raccoon with a strong S–S* expectancy of food may rub and dip and wash because these are its species-specific food-getting responses. The frightened rat may freeze rather than press the bar because this is one of its species-specific defense reactions. An animal may also have a previously learned R–S* expectancy that is stronger than that which the experimentally imposed reinforcement contingency can produce. In this case, too, we will find intruding behavior. Intruding behavior should occur particularly in the case of delayed opportunity for effective behavior. In other words, if reinforcement (S*) is delayed in the presence of cues (S) that signal reinforcement, we can expect intruding or adjunctive behavior to predominate.

Laws of Motivation

At this point, the cognitive model is essentially complete. It is complete with respect to its postulated structural determinants of behavior. But the model builder has the option here, as always, of introducing motivational concepts because he believes either that such concepts are convenient or because they are necessary. Motivational concepts are introduced here for the sake of convenience and in the hope that they will increase the model's flexibility and scope. The probability of a response increases with (*a*) the strength of the S–S* expectancy, (*b*) the strength of the R–S* expectancy, and (*c*) the value of S*.

There is a long tradition of regarding environmental cues as the source of incentive motivation, and the concept of S–S* expectancy follows this tradition. The strength of the R–S* expectancy is not a customary motivational factor; it merely determines the direction that motivated behavior will take—in this sense it is similar to Hull's construct of habit strength. The concept of value is submitted here in very rough and heuristic form. The value of S* may be positive, as food is when the animal is hungry, or negative, as when the animal expects shock. In the latter case, the effect of such an expectancy may be to inhibit behavior rather than motivate it. The value of S* certainly must depend on the physiological state of the organism; thus it has some of the conceptual properties of Hull's *D*. It should be emphasized, however, that while a procedure such as increasing deprivation may increase the value of S*, it does not necessarily increase response strength. According to this model, stronger behavior will be found only if the animal has, first, appropriate S–S* expectancies, and second, appropriate R–S* expectancies in the test situation. It is apparent that the most effective forms of reinforcement are precisely those S*s which are biologically significant or of value to the animal, such as food for the hungry rat. The extent to which an S–S* expectancy can be acquired when S* is insignificant, or lacking in value, remains to be determined, but the law of motivation indicates that even if such expectancies can be acquired, they will have minimal effect upon an animal's behavior. Roughly, then, the model suggests that the value of S* has some of the functional properties of Hull's *D*, the S–S* expectancy has some of the functional properties of Hull's *K*, and the R–S* has some correspondence with $_SH_R$. One might multiply them together to predict the strength of R.

I would deny that a direct associative linkage between a stimulus and a response is an important or interesting part of what is learned in most learning experiments. There are, of course, fixed S–R patterns that may be simple, like reflexes, or more complex, like releasers and fixed-action patterns. Also, S–R correlations that develop with sufficient training are probably learned. Such a mechanism would account for what is called automatic behavior (Kimble & Perlmuter, 1970), what is called ballistic behavior (Bindra, 1969), or what William James called habit. It would also account for well-learned behavior, motor skills, shifts in the associative control of behavior with continued training, and so on. A reinforcement process might be invoked at this point.[6] But perhaps sheer repetition of a response as a consequence of the law of performance suffices to connect it with prevailing stimuli. Certainly there is little a priori reason to expect such behavior to be governed by the same laws or to depend on the same neural mechanisms as those involved in the laws of learning, performance, and motivation that have just been proposed.

IMPLICATIONS

The present account of learning maintains that if an animal is placed in a situation where there are cues predicting food, and food is made contingent upon some response, the animal will learn, first, that these cues predict food and, second, that its behavior produces food. If the animal is hungry, then it is likely to make that response. This constitutes an explanation of why the reinforcement procedure works. This model also accounts for nonreinforced behavior, that is, the adjunctive behavior that occurs in superstition experiments, the instinctive drift phenomenon reported by Breland and Breland, and the ritualized, stereotyped behavior commonly found with a variety of reinforcement schedules. All of these phenomena follow from the assumption that an animal comes into the experimental situation with prior R–S expectancies and may or may not learn the particular R–S* expectancies it is supposed to learn. If the R–S* expectancy is learned, then the reinforcement contingency will provide effective control over the behavior, but not otherwise. While such a tactic for the explanation of behavior may appear to be circular, it is no more so than the more familiar tactic of assuming that a response is reinforced if it is strengthened, but not otherwise. In addition, it suggests how to start constructing rules that describe when the reinforcement procedure will work and when it will not.

Bindra (1972) has anticipated the first part of the present law of learning. His position is that learning consists only of the acquisition of S–S* representations. The cue is assumed to evoke a central motive state which, together with environmental supporting stimuli, dictates what response will occur. Bindra's analysis provides no counterpart to the present R–S* expectancy; the animal is restricted to making certain responses, such as approaching, contacting, and consuming the S* object. If the animal acquires other kinds of adaptive behavior, it has to be because there are intermediating cues which produce approach or contact responses. While this approach has parsimony, it seems likely to be costly in the ad hoc assumption of cues. The introduction of R–S* expectancies puts us a little closer to observable behavior, particularly for those kinds of animals that show great response flexibility. Flexibility is characteristic of much mammalian behavior (although defensive behavior in rats seems to be seriously constrained to innate behavior patterns; e.g., Bolles, 1970). I believe some mechanism is needed to account for this contrast, as well as the contrast in flexibility between birds' and mammals' behavior. Bindra's analysis appears appropriate for the learning of birds, which consists predominantly of responding in old ways to new stimuli. But to account for the plasticity of mammalian behavior, it may be useful to hypothesize the possibility of both stimulus (S–S*) and response (R–S*) learning.

Having looked at the implications for appetitive behavior, let us turn to one of the most difficult phenomena to explain—avoidance learning. I assume that invariably the initial learning consists of an S–S* expectancy in which the stimuli are the cues to shock (either an explicit CS or else generalized situational cues) and the shock itself. This expenctancy is evidently acquired very rapidly, often after a single shock; it has motivational properties and may perhaps be identified with what so many workers in the area have called fear.

---

[6] It would be ironic if the laws of reinforcement that have been so carefully worked out and vigorously defended in the effort to explain all instrumental learning should prove useful only for this kind of automatic behavior, that is, behavior that occurs when the organism is no longer perceiving the changing world around it, responding on the basis of environmental contingencies, or being motivated to achieve particular ends.

What happens next depends on what response the experimenter requires of the animal to avoid shock. If a response is required which is a strong species-specific defense reaction, for example, running away in the case of the rat, then the problem is virtually solved because the rat will start making this response almost immediately. One might think of this defensive reaction as being *elicited* by fear, and it may well be, in part. But it is also apparent that running away and freezing are in part, operant, or voluntary, or intentional kinds of behavior. That is, these defensive behaviors can be modified both by the cues in the situation and by their outcomes in the situation. They are, in short, subject to the laws of learning, at least in part, and should be explained, therefore, by the same paradigm that the law of performance applies to other learned behaviors. The paradigm requires a slight change, however, in the "psychological syllogism": The rat has a learned S–S* expectancy, cues predict shock, and an *innate* R–S* expectancy, *running away predicts safety*. The animal therefore runs away. The species-specific defense reactions can be regarded as innate R–S* expectancies in which certain responses innately predict safety. If the experimental situation is arranged so that such an expectancy is confirmed, then the problem is solved. But if the situation is arranged so that running away is impossible or fails to avoid shock, then this initially strong expectancy will soon be replaced by the more realistic one that running predicts shock; therefore, running will not occur. At this point the animal's behavior will be determined by the next strongest predictor of safety, another species-specific defensive reaction such as freezing.

What if none of the animal's innate defensive reactions are permitted to avoid shock; what if a rat is required to make an "arbitrary operant" like pressing a bar? The answer is empirically well established: the rat is unlikely to learn quickly or perform well, and it may evidence no learning at all. What is learned in this case is that all cues predict shock and all responses which the rat has any reason to make also predict shock. But if the animal begins to discriminate from the array of other events in the avoidance situation the contingency between the arbitrary response and safety, then as this realistic R–S* expectancy begins to replace the inappropriate, innate expectancy, bar pressing will begin to replace freezing and attempts to escape. But this is evidently a difficult discrimination. The rat's best chance of ultimately responding appropriately seems to rest upon learning to freeze at the bar, so that it can reflexively lurch at it when shock comes on which will minimize the negative value of S* (Bolles & McGillis, 1968).

The passive avoidance (punishment) situation is similar. The main difference is that in the case of punishment, the response usually has appreciable strength because the animal is busily doing something else for some reason. Because the response occurs and is followed by shock, the animal has a very good chance to learn not only about the imposed S–S* contingency and to make an innate defensive response, such as withdrawal, but also to learn the R–S* contingency. Then since the S* (the shock) has negative value, the punished response will be inhibited in accordance with the law of motivation. In summary, then, in the early stages of learning with aversive contingencies, if the response occurs because it is initially strong, then the animal will learn both an S–S* expectancy and perhaps an R–S* expectancy. In a case where the response contingency is reversed, as in an avoidance experiment, then the animal is only likely to learn an S–S* expectancy, and its behavior will be based on innate R–S* expectancies of safety.

The present model appears to account, at least in principle, for instances in which a positive reinforcement contingency controls behavior as well as the anomalous instances in which it does not. It deals, at least in outline, with what have always been more troublesome cases—punishment and avoidance acquisition. The model incorporates some relatively uninteresting aspects that can be called primary motivation (the value of S*) and some more interesting aspects that can be called learned or incentive motivation (S–S* expectancies). A recent review of incentive motivation (Bolles & Moot, 1972) indicates how the concept of predictive cues puts one immediately into contact with recent developments in such diverse areas as discrimination learning, secondary reinforcement, frustration, punishment, suppression, and avoidance. The model suggests easy extensions to decision-making situations in which subjects must select or make a choice between alternative cues that predict either different S*s or the same S* with different probabilities. Thus, we can readily translate from S–S* expectancy and S* value to belief strength and payoff utility.

The model makes no fundamental distinctions between Pavlovian and operant reinforcement situations beyond the simple observation that both of these traditional experimental procedures involve S–S* contingencies while only the latter adds the possibility for the subject to learn an R–S* expectancy. The model is based on the premise that animals learn expectancies that correspond more or less accurately to prevailing environmental contingencies. But it calls attention to the fact that when different animals are subjected to the same environmental contingencies, they may show quite different behaviors as a result of their different prior R–S* expectancies.

The model is quite deficient in one respect. It tells us very little about the practical application of the primary law of performance. It does not tell how quickly or how likely an S–S* expectancy and an R–S* expectancy will fit together so that a particular behavior will occur in a particular situation. The secondary law of learning does not even tell us when or under what conditions an R–S* expectancy will be learned. Moreover, most of the details of the model remain to be worked out; operational definitions of the theoretical terms need to be established; more than anything else, perhaps, the model needs to be quantified so that what are now heuristic statements can be made predictive.

Not all of the ambiguity can be attributed to the model, however; some of it is intrinsic in the phenomenon of learning. Our subject matter is enormously complex. In our own areas of specialization we can all think of a multitude of restraints upon what a particular subject can learn, and a multitude of factors that can prevent an expectancy from matching perfectly with the corresponding contingency. A general rule for how accurately expectancies match contingencies, therefore, must ultimately implicate the entire subject matter of psychology. A particular animal species, or a particular individual within a species, may readily learn one expectancy and be quite unprepared to learn another (Seligman, 1970; Shettleworth, 1972). The phylogenetic introduction of language and symbolization, the development of different stages of cognitive representation, the different sensory, motor, and motivational systems of different individuals all make it impossible to lay out any simple overall rules describing how fast expectancies will be formed or how readily they will be synthesized. Such difficulties are inherent. They constitute the subject matter of psychology. Is it not better to face them than to dismiss them with a facile reference to a universal law of reinforcement?

REFERENCES

BINDRA, D. The interrelated mechanisms of reinforcement and motivation, and the nature of their influence on response. *Nebraska Symposium on Motivation*, 1969, **17**, 1–37.

BINDRA, D. A unified account of classical conditioning and operant training. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning*. New York: Appleton-Century-Crofts, 1972.

BOLLES, R. C. Species-specific defense reactions and avoidance learning. *Psychological Review*, 1970, **77**, 32–48.

BOLLES, R. C. The avoidance learning problem. In G. Bower (Ed.), *The psychology of learning and motivation*. Vol. 6. New York: Academic Press, 1972.

BOLLES, R. C., & MCGILLIS, D. B. The nonoperant nature of the bar-press escape response. *Psychonomic Science*, 1968, **11**, 261–262.

BOLLES, R. C., & MOOT, S. A. Derived motives. *Annual Review of Psychology*, 1972, **23**, 51–72.

BOLLES, R. C., MOOT, S. A., & GROSSEN, N. E. The extinction of shuttlebox avoidance. *Learning and Motivation*, 1971, **2**, 324–333.

BOLLES, R. C., STOKES, L. W., & YOUNGER, M. S. Does CS termination reinforce avoidance behavior? *Journal of Comparative and Physiological Psychology*, 1966, **62**, 201–207.

Breland, K., & Breland, M. The misbehavior of animals. *American Psychologist,* 1961, **16**, 681–684.

Brown, P. L., & Jenkins, H. M. Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior,* 1968, **11**, 1–8.

Estes, W. K. Reinforcement in human learning. In J. T. Tapp (Ed.), *Reinforcement and behavior.* New York: Academic Press, 1969.

Falk, J. L. Conditions producing psychogenic polydipsia in animals. *Annals of the New York Academy of Sciences,* 1969, **157**, 569–593.

Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.

Hebb, D. O. On the nature of fear. *Psychological Review,* 1946, **53**, 259–276.

Hull, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review,* 1931, **38**, 487–506.

Hull, C. L. *A behavior system.* New Haven: Yale University Press, 1952.

Irwin, F. W. *Intentional behavior and motivation.* New York: Lippincott, 1971.

Katzev, R. Extinguishing avoidance responses as a function of delayed warning signal termination. *Journal of Experimental Psychology,* 1967, **75**, 339–344.

Kimble, G. A., & Perlmuter, L. C. The problem of volition. *Psychological Review,* 1970, **77**, 361–384.

Koch, S. Clark L. Hull. In W. K. Estes et al. (Eds.), *Modern learning theory.* New York: Appleton-Century-Crofts, 1954.

Logan, F. A. *Fundamentals of learning and motivation.* Dubuque, Iowa: Brown, 1969.

Moore, B. R. On directed respondents. Unpublished doctoral dissertation, Stanford University, 1971.

Mowrer, O. H. *Learning theory and behavior.* New York: Wiley, 1960.

Neuringer, A. J. Superstitious key pecking after three peck-produced reinforcements. *Journal of the Experimental Analysis of Behavior,* 1970, **13**, 127–134.

Rescorla, R. A., & LoLordo, V. M. Inhibition of avoidance behavior. *Journal of Comparative and Physiological Psychology,* 1965, **59**, 406–412.

Rescorla, R. A., & Solomon, R. L. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review,* 1967, **74**, 151–182.

Seligman, M. E. P. On the generality of the laws of learning. *Psychological Review,* 1970, **77**, 406–418.

Shettleworth, S. J. Constraints on learning. In D. S. Lehrman, R. A. Hinde, & E. Shaw (Eds.), *Advances in the study of behavior.* Vol. 4. New York: Academic Press, 1972.

Skinner, B. F. "Superstition" in the pigeon. *Journal of Experimental Psychology,* 1948, **38**, 168–172.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale University Press, 1956.

Staddon, J. E. R., & Simmelhag, V. L. The "superstition" experiment. A re-examination of its implications for the principles of adaptive behavior. *Psychological Review,* 1971, **78**, 3–43.

Tolman, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts, 1932.

Tolman, E. C. There is more than one kind of learning. *Psychological Review,* 1949, **56**, 144–155.

Trapold, M. A., & Overmier, J. B. The second process in instrumental learning. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning.* New York: Academic Press, 1972.

Walker, E. L. Reinforcement—"The one ring." In J. T. Tapp (Ed.), *Reinforcement and behavior.* New York: Academic Press, 1969.

Williams, D. R., & Williams, H. Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior,* 1969, **12**, 511–520.

(Received October 10, 1971; revision received April 18, 1972)

Psychological Review
1972, Vol. 79, No. 5, 410-419

PROCESSING DIMENSIONAL STIMULI: A NOTE[1]

G. R. LOCKHEAD[2]

Duke University

Analysis of available data shows that the nature of stimulus combinations, the particular stimulus dimensions used, affects how stimuli are perceived. Theories of stimulus processing, for example, serial or parallel, depend on how stimuli are perceived and thus depend on stimulus combinations. The data cannot be accounted for by any such singular processing model suggested to date. An alternative mode of processing which is in agreement with the data is proposed. The hypothesis suggests that a stimulus is first processed holistically, as a blob, which is defined. Subsequent processing of the blob into its components, or combinations of judgments of two or more blobs, occurs in a subsequent stage if the task requires. For object identification tasks, no later stages of processing are required.

There is current interest in the nature of the relations between dimensions which compose a stimulus; for example, are the stimuli analyzable, integral, or nonintegral? There is also considerable interest in how attributes of stimuli are processed, for example, in serial or in parallel. This article briefly summarizes these two approaches and then analyzes some of the studies directed at understanding how people process multidimensional stimuli; there is no intent to review the several hundred relevant studies in the literature. The analysis shows that there are available data which cannot be accounted for by any singular processing model; a parallel model and a serial model are both inadequate. An alternative which proposes that objects are initially processed by a holistic process is suggested. Then, and only if the particular task requires, the attributes of the stimulus are further analyzed and processed. It will be suggested that this second (and third or more) stage is conducted serially although this is not a logical requirement. More than one stage is a requirement.

Other authors have suggested two-stage processing (e.g., Bamber, 1969; Marcel, 1970; Neisser, 1967; Sternberg, 1967). This article is concerned with specifying the processing involved in the first stage of a model and with demonstrating the requirement of at least a second stage. Since I do not propose a complete model, one which specifies the aspects of the necessary second (or later) stages, other multistage models are not compared.

Relations between Dimensions

Integral stimuli. Shepard's (1964) paper is the first clear expression of concern with the form of the psychological metric as a function of the *particular* stimulus dimensions. He argued, from his work and from the developing work of Torgerson (1958) and Attneave (1950), that the similarity space which results from the combination of some dimensions produces a Euclidean metric, while other dimensions produce a different, perhaps city-block, metric. In each case the observers judged the similarity of stimuli. From this, Shepard suggested that there are two types of stimuli, "those that are related as homogeneous, unitary wholes, and those that tend to be analyzed into perceptually distinct components or properties [p. 80]." The first he termed "unitary" and the second "analyzable."

---

[1] This study was supported by Grant MH-18617 from the National Institute of Mental Health to Duke University. The author is grateful to the psychology departments at Stanford University and at the University of California, Berkeley, for providing the facilities at which this paper was prepared, and to Irving Biederman and Edward E. Smith for critical comments on the manuscript.

[2] Requests for reprints should be sent to G. R. Lockhead, Department of Psychology, Duke University, Durham, North Carolina 27706.

Data from a different experimental situation, absolute judgments of multidimensional stimuli, demonstrate that there can be a redundancy gain—performance improvement—when information is redundantly added to univariate stimuli (Eriksen & Hake, 1955). Later work showed that these redundancy gains are independent of receptor limitations; they occur both when the stimuli are easily seen and when they are made very difficult to see. This observation allowed the argument, analogous to the descriptions of Shepard, that there are multiple stimulus dimensions on which the observer can operate simultaneously. The term integral was devised to describe such stimuli (Lockhead, 1966a).

One reason for the distinction between the integral–nonintegral dichotomy and the unitary–analyzable dichotomy is to focus on the observation that "the problem is not that some stimuli are analyzable while others are not. . . . Almost certainly *S*s [subjects] can analyze any set of multidimensional stimuli [Lockhead, 1966a, pp. 103–104]." Shepard (1964) himself did not explicitly require that unitary stimuli could not be analyzed, and he allowed that "Of course, instead of a simple dichotomy, here, there may actually exist a graded series (from unitary to analyzable) [p. 80]." This observation is theoretically important, and there may be no stimuli which are unitary in the sense that the dimensions cannot be analyzed. Hue cannot be analyzed into its separate components and thus may be unitary. But hues cannot exist alone, and colors—stimuli which vary in hue, saturation, and brightness—are analyzable into dimensions and are also integral.

It is difficult to give an a priori definition of an integral stimulus. Phenomenologically, a Christmas tree light is a good example; we perceive its size, brightness, hue, and shape together, and it is difficult to perceive one of these aspects without the others. Recent work has further defined the concept of integrality by converging on that notion with different operations (Garner, Hake, & Eriksen, 1956). Garner and Felfoldy (1970) operationally define integral dimensions as

> those which lead to a Euclidean metric in direct distance scaling, produce a redundancy gain when the dimensions are correlated and some measure of speed or accuracy of discrimination is used, and produce interference in speed of classification when selective attention is required with orthogonal stimulus dimensions [p. 238].

This is the definition of integral stimuli assumed in this article.

Nonintegral stimuli. Stimuli "that tend to be analyzed into perceptually distinct components or properties [Shepard, 1964, p. 80]" are called nonintegral stimuli or separate stimuli here and are precisely what Shepard appears to have meant by analyzable stimuli. Operationally, these are stimuli which produce a city-block metric in direct distance scaling; there is no effect of the redundant separate dimensions on performance in speed tasks (Garner & Felfoldy, 1970), unless the stimulus elements are made difficult to see (Flowers & Garner, 1971); and there is an accuracy redundancy gain in case of process limitation (Lockhead, 1970). A two-dimensional nonintegral stimulus is considered here to be two integral stimuli, as if it were two things.

Analyzable stimuli. Integral and nonintegral stimuli are both analyzable. Operationally, analyzable stimuli are composed of dimensions for which orthogonal variation on one dimension does not affect identification performance on the other dimension. Egeth and Pachella's (1969) data demonstrate that univariate identification judgments of a stimulus can be the same when the stimulus varies along only that dimension as when it varies along two dimensions simultaneously. Thus the stimulus, a dot which could vary along either the horizontal or the vertical, is analyzable. This appears also to be an integral stimulus since, as shown by Klemmer and Frick (1953), the same amount of information is transmitted by a single dot which varies in two dimensions as by two different dots which vary in one dimension.

Processing Models

Many different models which contrast serial with parallel processing have been proposed to account for the processing of more than one stimulus aspect or for the performance of multiple tasks. Recent considerations of some of these several models have been offered by Biederman and Checkosky (1970), Egeth (1966), Hawkins (1969), Neisser (1967), Nickerson (1970), Posner (1964), Rapaport (1959), Smith (1968), Sternberg (1969), and others, most of whom are referenced in these studies. All of the single-stage models which have been proposed require independence (or less than perfect correlation), at some stage, of two or more dimensions, and Hawkins (1969) has summarized the behavioral requirements of nine of these. Simply, the various assumptions require that aspects are processed either serially or in parallel and that this processing can be either exhaustive or self-terminating. The order of searching a dimension can be ordered or random. In all cases, decisions made on one dimension do not affect decisions on other dimensions although, in some models, the presence of a second dimension may attenuate or otherwise affect discriminability of the first dimension (cf. Lindsay, 1970). Empirical attempts to discriminate between these several differing possibilities have produced only negative or equivocal results; it may be that neither parallel nor serial processing is an appropriate analysis.

Some Interactions between Dimensional Relations and Processing Model Predictions

There are data in the literature which must be addressed both by models of stimulus combination and by models of processing. Three such results are important for this concern with the way stimuli are processed and are pointed to here.

Stimuli do not have to be integral to provide an accuracy of discrimination redundancy gain. If errors are being made in an absolute-judgment task, identification performance improves with increased dimensionality of the stimuli. This is as long as the added dimensions do not block out or otherwise interfere with the perceptibility of the present dimension. This statement holds even though performance accuracy on any one dimension may decrease when the number of dimensions is increased (cf. Egeth & Pachella, 1969; Garner, 1962; Lockhead, 1966a; Miller, 1956) and holds when the dimensions are integral (e.g., Eriksen & Hake, 1955; Lockhead, 1966a) or nonintegral (Lockhead, 1966b, 1970). In an absolute-judgment task, redundant combinations of dimensions in different modalities—lightness with loudness and roughness with hue—can provide large redundancy gains (Lockhead, 1970). These stimuli do not conform to the phenomenological definition of integral, and they probably do not conform to the operational definition either. Almost certainly there will be no interference in a speeded task which requires selective attention when finger roughness is added to hue. Consideration of only this class of results might suggest that the nature of the stimulus dimensions is not important to processing models.

Selective attention. A filtering task (Posner, 1964) requires decisions based on one dimension, while a second and irrelevant dimension also varies. Garner and Felfoldy (1970) clearly demonstrated that for integral stimuli, sorting speed along one dimension is faster when a second dimension does not vary than when it does. The second dimension produces interference. But if this same second dimension varies in a correlated fashion with the first dimension, there is no interference. In fact, there is facilitation. That is, there is no interference in the processing of the entire stimulus which can be attributed to the presence of the second integral dimension, but there is interference when the task is to attend selectively, when the observer is asked to analyze or otherwise disentangle one dimension from the complex. It is important to note that the filtering task is performed accurately even though sorting speed is affected. This means that the integral stimuli are analyzable.

When stimuli are nonintegral, which can be achieved by physically separating the stimulus dimensions into different patches or things, Garner and Felfoldy's data show that there is no effect of adding a second dimension. There is now nothing to be disentangled. Contrary to the previous section, these results suggest that the nature of the stimulus dimensions is critically important to the processing model. They also suggest that the effects of the stimulus dimensions on processing interact with the experimental task.

The form of redundancy. In most performance tasks designed to determine redundancy gains, the experimenter chooses two stimulus dimensions which appear to be psychologically independent and continuous, such as hue, lightness, angle, and extent. Some number, *m*, of (usually) equally spaced values from each dimension is judged separately in an absolute-judgment task. The *m* linearly correlated two-dimensional stimuli are also judged. If observers can identify the two-dimension linearly correlated set with greater accuracy than they can identify either of the single-dimension sets, there is a redundancy gain.

Warren Torgerson (personal communication, 1965) was the first to ask why the single dimension sets were paired in a linear correlation rather than some other correlation such as the sawtooth correlation indicated by the Os in Figure 1. Since then, several "naive" undergraduate students have asked the same question, but without a descriptive term like sawtooth. Observe that, in Figure 1, the X stimulus set and the O set have the same uncertainty. For both sets, each value of one dimension is paired with one and only one value of the other dimension. However, although the uncertainties are equivalent, the discriminabilities are not. Stimuli which are sawtooth paired (the Os) are identified with considerably more accuracy than are stimuli which are linearly paired (Lockhead, 1970). The data in that paper did not include reaction times, thus processing models which use response times as the critical dependent variable cannot be evaluated directly with those data. It is logically possible that observers use different processes when attempting to identify accurately two-dimensional stimuli than when attempting to respond as quickly as possible to them.

Levy and Haggbloom (1971) did collect reaction times in this task and their data are of importance here. Using size of square and brightness of patch as the two dimensions, they obtained reaction times on the conditions of interest, the set of Xs and the set of Os of Figure 1. Although they concluded from the study that there were no significant differences between redundant sets, the conclusion is too conservative. Levy and Haggbloom found a significant interaction between practice and reaction times to the correlated sets. They suggested, as seems reasonable, that the O set may have required longer times early in training because of response compatibility differences and they referenced Fitts and Seeger (1953) who demonstrated that natural orderings between stimuli and responses produce faster reaction times than do other orderings. Visual inspection of their data shows that by the third experimental session, reaction times are reliably faster to the set of O than the set of X stimuli. This result has since been replicated with the conclusion that the sawtooth-related stimuli are reliably identified faster than are the linearly related stimuli; and both correlated sets are identified faster than either univariate set (Levy, personal communication, 1971).

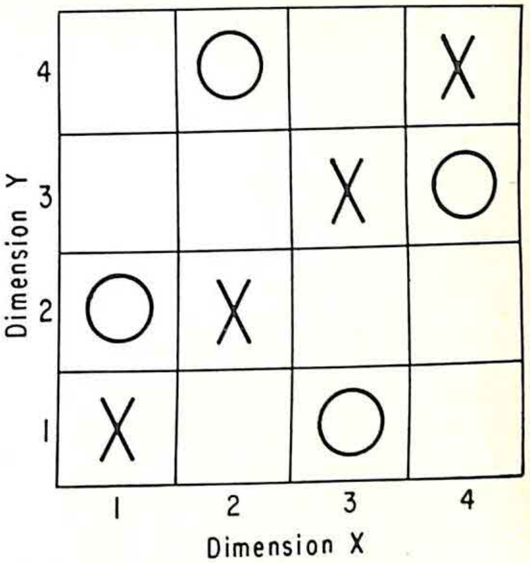


FIG. 1. The stimulus pairings considered in the analysis of the integral space. (The Xs represent the linearly paired set; the Os represent the sawtooth-paired set. Each dimension is any ordered psychological dimension.)

Levy and Haggbloom's Result and Processing Models

It is useful to carefully analyze processing models in terms of the finding that reaction times are faster with the sawtooth-related than with the linearly related integral stimulus set. The most powerful of the several suggested processing models appears to be one which assumes self-terminating parallel processing with variable and independent processing times along each dimension as proposed by Hawkins (1969) and independently by Biederman and Checkosky (1970). Other singular processing models which have been proposed have been rejected by the demands of the data considered in those papers or they have more constraints than does this one. In the latter case, those models and this one predict the same results (cf. the concise arguments given by Hawkins, 1969). (The added possibility that the processing rate of one dimension is attenuated by the presence of a second dimension [Triesman & Geffen, 1967] or that processing rates covary [Hawkins, 1969, Assumption 4] is not discussed. The comparisons to be made here are between two-dimensional tasks, and any effects of the added dimension should be the same in both tasks.)

Such parallel-processing models can be outlined as a two-horse horse race. Each horse is a value on one dimension and is processed independently of the other horse. Processing means that the organism decides that a particular stimulus value of a dimension is (or is not) the criterion value. As soon as a decision is made, on either channel, a response is accessed. Whichever horse reaches the finish line first ends the race. (Experimentally, we are considering a correlated task, one in which cell x_1y_1 means press the left finger and cell x_2y_2 means press the right finger, and the other two cells in the potential 2×2 stimulus matrix do not occur.) The observer can respond on the basis of his judgment of the value of either dimension X or dimension Y; x and y are independent of each other and are both distributed in some manner over time. Thus, according to the race horse model, on those trials that x is faster, x crosses the finish line first. When y is faster, he stops the race. When either x or y could stop the race, the correlated case being considered here, the expected distribution of times to the end has a lower mean than either of the components. This is because the probability that one of the two horses will have, on any given trial, a response time of less than each mean (when the component distributions have equal means) is .75. Reaction times are, in fact, faster in the correlated case than in the univariate case. This model of processing times is similar to other models of processing accuracy (cf. esp., Hake, Rodwan, & Weintraub, 1966). The parallel-processing model "provides a parsimonious account of the difference in RTs [reaction times] between the information-reduction tasks defined by Posner's (1964) taxonomy [Biederman & Checkosky, 1970, p. 490]." The authors' data allow rejection of a strict serial-processing model for the task involved and reliably predict reaction times when discriminability of the dimensions is separately manipulated. They do not extend their predictions to experimental procedures which employ more than two values on a dimension.

Going now beyond that procedure, consider the prediction of a horse race model for the set of four O stimuli as depicted in Figure 1. (The important aspect of the following analysis is that there must be three or more values on a dimension. Most studies concerned with processing models use only two values on each of two dimensions and the point important to this analysis is confounded.) There are four stimuli and four responses and each response can be determined by its value on the X dimension or on the Y dimension or on both. According to the horse race model, responses should be faster, on the average, to the O set and to the X set than to either of the unidimensional sets (four stimuli which vary only in X or which vary only in Y). This is for the same reasons as outlined above. With the correlated set, there are two opportunities (or horses) to provide a shorter latency for each of the four stimulus–response pairings. There is no reason, according to that model (or, to my knowledge, according to any other single-stage model suggested), to consider that reaction times to the X set should be any different than reaction times to the O set. The same element decisions are being made in each case. However, reaction times are faster to the O set than to the X set. The stimuli were integral.

Levy and Haggbloom's (1971) result is inconsistent with any model proposed to date which requires independent decisions on the separate dimensions. For such models to account for the sawtooth results, it is necessary that one independent processor know something about the decision(s) of the other processor, that there be some kind of communication across processors. One could add a new process to the parallel model which would utilize incomplete processing on each dimension. Such a process would evaluate Dimension X and then, because of the nature of the stimulus pairing, change its criterion for Dimension Y based on the partial information from x_i. Then the partial and conditional information from y_j would affect the next decision on X. This multiplexing process would continue until some overall criterion was met after which a response would be effected. Such a model could be developed to account for the finding that sawtooth-paired integral stimuli are identified faster than are linearly paired stimuli.

But it appears that even such a complex model would not predict the results of a filtering task with integral stimuli, the observation that the trial-to-trial varying of a second dimension interferes with performance on the first dimension. Any information and any interference from the second dimension, in such a model, ought to be the same in a filtering task as in a univariate task. Rather than make further modifications to account for the data with a parallel model, a slightly different mode of processing is suggested as a working hypothesis. The hypothesis has the advantages of being consistent with all of the data considered and of requiring fewer assumptions than the models discussed.

Blob Processing

The alternative suggested here is that a stimulus is first processed or perceived holistically; the resulting holistic configuration is called a blob. If the task can be performed on the basis of just this processing, as when identifying an integral stimulus, a response is accessed. If not, blob processing is followed, as necessary, by serial processes. This is a dynamic model in the sense that the organism employs only those processes necessary for the task. If serial processing is not required by the particular stimulus dimensions and task selected by the experimenter, it is not used. Blob processing is always employed. The suggested model is described below by illustration of the resulting processing of different stimuli in different tasks.

Processing integral stimuli. Consider two (or more) dimensions which are continua in the same Euclidean psychological space, a requirement for integrality. To identify a particular stimulus, assuming a practiced observer, the "*S*'s [subject's] task is one of locating the multidimensional stimulus in that space without the requirement that the space be analyzed according to its separate dimensions . . . [Lockhead, 1970, pp. 2–3]." This is the blob. According to this notion, the "*S*s [subjects] learn to attach a response label to a location in a psychological space rather than learn the combination of the separate elements [Lockhead, 1970, p. 10]." In a discrimination space such as this, the mean distance between any O stimulus (cf. Figure 1) and its immediate neighbors is greater than is the mean distance between any X figure and its neighbors. Thus, sawtooth-paired stimuli are easier to discriminate from each other than are linearly paired stimuli; responses are more accurate (Lockhead, 1970) and are faster (Levy & Haggbloom, 1971) with sawtooth-related than with linearly related stimuli. The horses related to the Os are faster even though the Os and the Xs have the same ancestors. This is not parallel processing; an integral stimulus is a single thing and there is no requirement for independent decisions on the separate dimensions.

When the task requires identification of the stimulus, correlated integral stimuli are identified at least as accurately and rapidly as are univariate stimuli. If there are identification errors with the univariate stimuli, integral stimuli are identified faster and more reliably than are univariate stimuli. This is because, according to the hypothesis suggested, both tasks require the same blob processing, and adjacent correlated stimuli are separated further in the discrimination space than are adjacent univariate stimuli.

Sometimes the task imposed by the experimenter requires integral stimuli to be judged in ways other than identifying a locus in a multidimensional space. In a filtering task, the observer has to judge the value of a single dimension while the stimuli vary on other dimensions as well as on that one being judged. According to a blob-processing notion, this requires, once the blob itself has been processed, that the perceived object is then analyzed into its component values. This is because the task is not to identify or recognize the thing presented; it is to identify (abstract) an aspect of it. Thus filtering tasks require more time than do identification tasks (Garner & Felfoldy, 1970). It takes less time to respond on the basis of all aspects of an integral stimulus than it does to respond on the basis of any one aspect of it.

Condensation tasks require even more time to perform than do filtering tasks (Fitts & Biederman, 1965). Since filtering tasks require at least two stages, perhaps condensation tasks require at least three stages. In a condensation task, the observer makes one response to the two stimuli x_1y_1 and x_2y_2, and a different response to the other two stimuli, x_1y_2 and x_2y_1. In the sense of the analysis of variance, the interaction terms are treated as the main effect, and this then becomes a filtering task. The observer must determine the separate value of a complex stimulus on *both* varied dimensions *and* he must make his response contingent on the combination of these two analytic determinations. This is something analogous to performing mental arithmetic and is performed after the second stage of the model. It may be that the blob is first processed, then the blob is sequentially analyzed to determine both component values, and finally the results of these sequential decisions are combined to determine a response. Such an analysis is consistent with the data.

In all of these instances, the observer processes the integral stimulus as a thing, as a blob, just as rapidly as he does any stimulus. In fact so-called univariate stimuli are single dimensional only in the sense that the experimenter is varying only one attribute from trial to trial. But the stimulus being perceived is multivalued. This obvious but essential point is made by considering what the observer would report in a situation that changes in ways not essential to the experimental question. For example, imagine a choice reaction time task which requires discrimination between orientations. The stimulus has been a 2-inch-long black line on a white surround. On one trial, the stimulus presented is unexpectedly a green line on a purple surround, or it could be dotted, or an array of lines, or any of a large number of changes which do not affect the attribute being judged. If the observer is judging the single dimension of orientation, these changes in the physical stimulus will go unnoticed. If the observer perceives a blob and then abstracts the appropriate information out of the blob, he may—and will—suggest that there is something wrong. Observers do not judge univariate stimuli; they make univariate judgments of multivariate stimuli. Observe that if dimensions such as line color vary regularly, this becomes a filtering task, and there is a continuing effect of changes in the irrelevant dimension.

How a blob would be processed in terms of physiology or function is not known or suggested. To point to the feasibility of something like blob processing, it is interesting to observe that the hypothesis is consistent with an analogy to complex cortical cells. Hubel and Wiesel (1965) show that there is an optimal stimulus for a single cell in Areas 18 and 19 of the visual cortex of cat. For that cell being recorded, the particular stimulus might be a line of a certain orientation and length which would occur at different sites (within about 5° of visual angle) on the retina. Any change in the length or in the orientation of that stimulus on the retina results in a decreased rate of firing of the complex cell. At this particular neural site, the stimulus is not analyzed into its components; the cell responds to the multidimensional integral stimulus. A different integral stimulus optimally stimulates a different complex cell. Ostensibly, there is a large number of such cells and, by this analogy, each may reflect a different blob. Certainly, then, the notion of blob processing does not violate any physiological evidence although it is yet too great a leap to talk of integrality and complex cells in the same manner.

Processing nonintegral stimuli. According to the experimenter's definition, a stimulus card which contains two separate stimulus patches may be a single stimulus. According to the observer, a nonintegral stimulus is two integral stimuli presented together. When stimuli are nonintegral, each aspect of the stimulus is a separate thing and *a two-dimensional nonintegral stimulus is two blobs.* (The two attributes do not have to be physically separated to be nonintegral. For example, two-dimensional stimuli which vary in size of circle and in the orientation of a radius drawn within the circle present all aspects of the stimulus in the same physical place and are nonintegral [Shepard, 1964].)

If the task to be performed can be satisfactorily accomplished by identification of only one of the two blobs in a nonintegral stimulus, there is no effect of the presence of the second blob. Thus, according to the model and the data, in a filtering task the added dimension does not interfere and in a speeded redundant task the added dimension does not facilitate (Garner & Felfoldy, 1970).

If the task is not speeded, and if there are limitations in the identification of any one dimension, redundant nonintegral information does facilitate (Lockhead, 1969b, 1970). Accuracy of identification of redundant nonintegral stimuli is facilitated as follows. The two separate blobs are identified independently, but because of process and/or state limitations (cf. Garner, 1970), identification accuracy on the basis of either of the blobs is limited. Thus, the noisy identifications of the separate blobs are combined in a later stage, just as mental algebra or geometry might be done, to produce a combined judgment of the experimenter's value of the single stimulus. According to this notion, reaction times should be as long or longer when identifying nonintegral multidimensional stimuli which produce an identification redundancy gain than when identifying univariate stimuli. That is, Levy and Haggbloom's results might have been different if the stimuli were nonintegral. This is not yet tested.

Processing complex stimuli. It is rare to encounter objects in the real world which are as simple as those used to investigate processing models. It is even more rare to encounter entire stimulus sets in which the objects differ from each other in only one or two ways. By many orders of magnitude, real things such as trees and cups and faces are each more complex, and the differences between such objects are more various, than anything considered here. The ideas expressed in this paper are not sufficient to account for the identification of these more complex things.

Yet these notions do lead to a conjecture. Most theories of pattern recognition concern (*a*) the analysis of stimulus attributes and (*b*) the subsequent combination of decisions concerning these attributes or distinctive features to provide an identification (cf. Neisser's 1967, Ch. 3, discussions of several models). One difficulty in interpreting such models is determining what is to be called a feature and which features are to be employed to achieve recognition. Another difficulty concerns the number of features needed. An object cannot be reliably identified on the basis of a particular feature if other objects have the same feature. Also, we can reliably identify only a few values on a particular dimension (there is channel capacity), and there are few simple stimulus attributes or dimensions (e.g., extent, brightness, orientation, etc.). Thus, it is extremely unlikely that there is *any* object which does not have attributes in common with many other objects. This means that we do not ordinarily identify objects on the basis of one or even several features. A distinctive feature must be a set of attributes considered in relation to all stimuli; one cannot have distinctive features in a vacuum.

That cartoonists may draw caricatures employing one particular feature to identify a person or thing does not imply that identification is accomplished by means of only that feature. Bob Hope's nose on any other person would not be Bob Hope. Alternatively, it is very likely that any other nose on Bob Hope would provide an object identified as Bob Hope with someone else's nose. Even more extreme, the removal of the arms from any statue in Trafalgar Square does not produce a new Venus DeMilo. Distinctive features are employed in a context. In Figure 1, if we add another stimulus which has an extremely large value on just one of the dimensions, that would be called a distinctive feature, but only because it was removed from other features in the set and not because of anything inherently distinctive in the feature itself. Probably no theorist has meant to suggest that only distinctive features are employed. However, the less obvious implication is not often stated. We must determine the space, the set of relations, and not just the features, if we are to understand pattern recognition.

The basic hypothesis is that observers first locate an object in some complex psychological space and then analyze that locus according to the needs of the task. (Unfortunately, how that locus is realized is taken as a given. But we first have to determine that there is such a set of relations before we can ask how they are determined.) If we do identify each familiar integral stimulus as if it were a locus in some identification space, perhaps a distinctive feature can be defined as an attribute(s), or the value of an attribute(s), of a stimulus which causes that integral object to be distant from other potential stimuli in the psychological space. This does not solve the original problem. It simply directs attention to the possibility that the relations between attributes (which is another way of saying locus in the space) may be processed *before* the values of the attributes themselves are processed.

One aspect of this notion is that the locus of the space is what is operated on when learning a new object. In a features model, only those aspects abstracted are logically available for later use. For the blob notion, all attributes in the stimulus are potentially available for later use (arbitrarily ignoring memorial difficulties, preferences, attention, and saliences previously learned or native to the observer). Thus complex objects which have not been processed frequently enough or long enough to be translated into some cognitive or verbal structure are nonetheless seldom confused with other complex objects, and once a complex scene has been viewed, that overall locus, and others near it, are recognized later even though details may not be known. A demonstration by Shepard (1967) points strongly to the viability of the suggestion that it is a locus in space which is coded. He showed that 612 different pictures are recognized with essentially no errors, in a forced-choice situation, following a delay of two hours since seeing each picture only once. To consider seriously that the observers abstracted and encoded 612 separate salient attributes, remembered them with essentially no error, and then knew which features to abstract from the test stimuli to compare with memory, is simply too implausible.

Conclusion

It is hypothesized that observers process integral stimuli as separate things and that the aspects of dimensions which compose a thing are not ordinarily evaluated separately. In tasks which cannot be performed satisfactorily without evaluation of aspects of the stimulus, the integral stimulus is then further analyzed into its components. Some tasks require still additional processing once the component values have been judged. The time required to process a stimulus is correlated with these requirements. If a stimulus can be responded to on the basis of the blob, responses are fast.

The model requires consideration both of the nature of the stimulus dimensions and of the experimental task in any evaluation of the mode of stimulus processing. It also suggests that initial processing is neither serial nor parallel, but that integral stimulus dimensions are first processed simultaneously in a blob and are as a single thing. There is every reason to suppose, as in reading this page, that processing of several integral stimuli—processing nonintegral stimuli, which is to say the same thing—is serial. There do not currently appear to be any data which require the proposal of parallel decision processes, although such a possibility has not been logically rejected.

REFERENCES

Attneave, F. Dimensions of similarity. *American Journal of Psychology*, 1950, **63**, 516–556.

Bamber, D. Reaction times and error rates for "same"–"different" judgments of multidimensional stimuli. *Perception and Psychophysics*, 1969, **6**, 169–174.

Biederman, I., & Checkosky, S. F. Processing redundant information. *Journal of Experimental Psychology*, 1970, **83**, 486–490.

Egeth, H. E. Parallel versus serial processes in multidimensional stimulus discrimination. *Perception and Psychophysics*, 1966, **1**, 245–252.

Egeth, H., & Pachella, R. Multidimensional stimulus identification. *Perception and Psychophysics*, 1969, **5**, 341–346.

Eriksen, C. W., & Hake, H. W. Multidimensional stimulus differences and accuracy of discrimination. *Journal of Experimental Psychology*, 1955, **50**, 153–160.

Fitts, P. M., & Biederman, I. S–R compatibility and information reduction. *Journal of Experimental Psychology*, 1965, **69**, 408–412.

Fitts, P. M., & Seeger, C. M. S–R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 1953, **46**, 199–210.

Flowers, J. H., & Garner, W. R. The effect of stimulus element redundancy on speed of discrimination as a function of state and process limitation. *Perception and Psychophysics*, 1971, **9**, 158–160.

Garner, W. R. *Uncertainty and structure as psychological concepts.* New York: Wiley, 1962.

Garner, W. R. The stimulus in information processing. *American Psychologist*, 1970, **25**, 350–358.

Garner, W. R., & Felfoldy, G. L. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1970, **1**, 225–241.

Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationism and the concept of perception. *Psychological Review*, 1956, **63**, 149–159.

Hake, H., Rodwan, A., & Weintraub, D. Noise reduction in perception. In K. R. Hammond (Ed.), *The psychology of Egon Brunswik.* New York: Holt, Rinehart & Winston, 1966.

Hawkins, H. L. Parallel processing in complex visual discrimination. *Perception and Psychophysics*, 1969, **5**, 56–64.

Hubel, D. H., & Wiesel, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 1965, **28**, 228–289.

Klemmer, E. T., & Frick, F. C. Assimilation of information from dot and matrix patterns. *Journal of Experimental Psychology*, 1953, **45**, 15–19.

Levy, R. M., & Haggbloom, S. J. Test of a multidimensional discrimination model of stimulus identification. *Psychonomic Science*, 1971, **25**, 203–204.

Lindsay, P. H. Multichannel processing in perception. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis.* New York: Appleton-Century-Crofts, 1970.

Lockhead, G. R. Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, 1966, **72**, 95–104. (a)

Lockhead, G. R. Visual discrimination and methods of presenting redundant stimuli. *Proceedings of the 74th Annual Convention of the American Psychological Association*, 1966, **1**, 67–68. (Summary) (b)

Lockhead, G. R. Identification and the form of multidimensional discrimination space. *Journal of Experimental Psychology*, 1970, **85**, 1–10.

Marcel, A. J. Some constraints on sequential and parallel processing and the limits of attention. In A. F. Sanders (Ed.), *Attention and performance III.* Amsterdam: North Holland Publishing, 1970.

Miller, G. A. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **65**, 81–97.

Neisser, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

Nickerson, R. S. *Binary-classification reaction time: A review of some studies on human information-processing capabilities.* (Rep. No. 2004) Cambridge, Mass.: Bolt Beranek & Newman, 1970.

Posner, M. I. Information reduction in the analysis of sequential tasks. *Psychological Review*, 1964, **71**, 491–504.

Rapaport, A. A study of disjunctive reaction times. *Behavioral Science*, 1959, **4**, 299–315.

Shepard, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, **1**, 54–87.

Shepard, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 156–163.

Smith, E. E. Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 1968, **69**, 77–110.

Sternberg, S. Two operations in character recognition: Some evidence from RT measurement. *Perception and Psychophysics*, 1967, **2**, 45–53.

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II.* Amsterdam: North Holland Publishing, 1969.

Torgerson, W. S. *Theory and methods of scaling.* New York: Wiley, 1958.

Triesman, A., & Geffen, G. Selective attention: Perception or response? *Quarterly Journal of Experimental Psychology*, 1967, **19**, 1–18.

(Received February 15, 1972)

Psychological Review
1972, Vol. 79, No. 5, 420–427

THEORETICAL NOTES

ON REASONING, CONGRUENCE, AND OTHER MATTERS[1]

JANELLEN HUTTENLOCHER[2] AND E. TORY HIGGINS

Teachers College, Columbia University

This paper deals with issues raised by H. H. Clark's rebuttal in 1971 of a 1971 paper by J. Huttenlocher and E. T. Higgins. The first section shows that findings with negative-equative sentences, because of their dependency both on experimental context and on the particular adjectives used in the problems, are of questionable value in evaluating Huttenlocher and Higgins' hypothesis of a parallel between syllogistic reasoning and tasks involving the arrangement of real items. The second section shows that Clark's "congruence" model is lacking, not only as an explanation for question answering in syllogisms, but, more important, as a general model for question answering and verification tasks; it is theoretically uneconomical and would seem to lead to predictions contrary to known evidence. Finally, it is shown that the evidence Clark cites in support of his principle of lexical marking is inadequate.

In this paper, we consider various points raised by Clark (1971) in his rebuttal to our paper "Adjectives, Comparatives, and Syllogisms" (Huttenlocher & Higgins, 1971).

On Reasoning

Huttenlocher (1968) presented behavioral evidence in support of subjects' claims that they solve ordering syllogisms by constructing imaginary spatial arrays of the items. In brief, subjects claim to imagine arranging the two items from the first premise and then to imagine "adding in" the new item from the second premise. The evidence consists of a correlation between the difficulty of adding in that new item from the second premise in syllogisms and the difficulty of adding a new *real* item into a fixed array of two other *real* items. For both, it is easier to arrange the new item when it is the grammatical subject of a regular comparative or of an active sentence, or the grammatical object of a passive sentence (using "is led" and "is trailed").

As Huttenlocher and Higgins (1971) pointed out, it is possible to explain the observed correlation without postulating that subjects actually treat syllogisms as imaginary arrangement tasks. The argument is that the reason for *both* findings is that the difficulty of understanding the relation between two items when one item is *new* and the other *old* depends on the grammatical roles assigned to those new and old items in a sentence. Such a formulation, in terms of "new" and "old" items makes the same behavioral predictions as the imagery formulation, but fails to account for subjects' introspections about how they solve such problems. Whichever formulation proves best, both explain the process of combining information from the two premises which, as Clark (1971) now admits, he has failed to do. Even if imagery does prove to be directly implicated in the solution process, one could not accurately describe our proposal as an "imagery theory" if, by that, it were meant that imagery was the sole basis for solving the problems. Clearly, the imagery generated in problem solving is the product of various logical and linguistic processes, nor did we ever claim otherwise.

Clark's (1971) rebuttal argues that for syllogisms with negative-equative premises, for example, "Sam is not as good as Joe," we must predict that problems should be easiest either when the new item is grammatical subject or when it is grammatical object. The former, he says, is the more "correct" proposal for us to make. This is a mistaken version of our proposal. We only predict a correlation between the difficulty of negative-equative syllogisms and the difficulty of arranging real mobile items with respect to real fixed items according to negative-equative instructions. We compared

---

[1] The preparation of this paper was supported in part by Career Development Award 5-K-4-HD-21,979, and in part by Research Grant HD 03215, both from the National Institutes of Health to the senior author. The authors thank Carol Milligan for her help in running the subjects in the experiment reported in this paper.

[2] Requests for reprints should be sent to Janellen Huttenlocher, Teachers College, Box 97, Columbia University, New York 10027.

TABLE 1

MEAN REACTION TIMES AND PERCENTAGE OF ERRORS FOR NEGATIVE-EQUATIVE ARRANGEMENT TASKS

Adjective	Subject mobile		Object mobile		Adjective	Subject mobile		Object mobile	
	RT	%	RT	%		RT	%	RT	%
good	1.76	7	1.82	7	bad	2.24	10	2.12	18
pretty	1.90	8	1.78	5	ugly*	2.42	18	2.13	16
happy	1.99	7	1.80	6	sad	2.18	15	1.97	14
big	1.84	9	2.04	8	small*	2.31	6	2.07	18
tall	1.94	1	1.90	8	short*	2.65	12	2.13	17
fast	1.79	7	1.57	8	slow*	2.50	12	1.96	24
light*	2.49	16	2.18	16	dark*	2.41	11	2.10	14
hot*	2.36	13	2.05	18	cold	2.54	25	2.31	12
rich	1.99	5	1.95	2	poor*	2.45	13	2.00	11

Note.—RT = reaction time. Errors and RTs for each adjective are from 60 subjects with two trials, that is, 120 problems.
* RT differences significant at $p < .01$, Wilcoxon signed-ranks test.

the difficulty of negative-equative syllogisms and negative-equative arrangement tasks (Huttenlocher, Higgins, Milligan, & Kauffman, 1970) and found the following: (*a*) For *syllogisms,* the object effect was highly significant for *bad,* whereas for *good* it was in the same direction but not significant. (Clark found significant differences for both *good* and *bad,* but differences with *good* were slight. In his results, as in ours, object effects with negative-equatives are proportionally smaller than subject effects with regular comparatives.) Huttenlocher et al. (1970) obtained no object effect with *high–low* syllogisms for either adjective; indeed, the results, although insignificant, were in the opposite direction. Clark completely ignores these findings and argues as if there were, in general, an object effect for negative-equative syllogisms. (*b*) For *arrangement tasks* using these same adjectives, there was again a significant object effect only for *bad,* not for *good.* For *high* and *low,* there was an insignificant subject effect. Our prediction of a correlation was thus confirmed even though there was *not* a universal subject or object effect with negative-equatives for either arrangement tasks or syllogisms.

Clark (1972a) reports a study where negative-equative arrangement tasks with *high, low,* and *bad, good* are all easier with the mobile item as subject rather than object. He claims this result as evidence against our proposal because of its striking contrast to results with negative-equative syllogisms. However, as we saw in *a* above, Clark misrepresents our results with negative-equative syllogisms by claiming there is always a significant object effect. Further, Clark's results for arrangement tasks differ markedly from ours, since the Huttenlocher et al. arrangement tasks showed an object effect for *bad* and no significant difference for *good, high,* and *low.* In fact, there is considerable evidence that experimental setting strongly affects results for negative-equative arrangement tasks. Clark argues that the reason for the discrepancy between his results and those of Huttenlocher et al. on negative-equative arrangement tasks is that the latter used the format of a syllogism which has a question at the end. However, Table 1 presents results from a negative-equative arrangement experiment which did *not* have a syllogistic format and, nevertheless, found once again object, not subject, effects. Using the apparatus from Huttenlocher et al., on each trial, the experimenter placed two items, describing the relation between them as he did so (randomly varying description form). The subject had to place a third item with respect to these two according to a negative-equative instruction given by the experimenter. Several insoluble problems were included. Negative-equative instructions involved 18 adjectives (9 polar pairs). Most subject–object differences were insignificant. The significant differences were all *object* effects, and, as predicted by Huttenlocher et al., most of the significant effects occurred with marked adjectives.

Since our proposal postulates a parallel between syllogisms and arrangement tasks, we must consider what constitutes a good test of whether such a parallel exists. This problem only arises with negative-equatives because they, as opposed to the other sentence forms we have used, produce variable results depending on the adjective used in the problems, and, for arrangement tasks, on at least some of the following experimental setting factors as well: (*a*) whether subjects have to arrange the first two items themselves, (*b*) whether the placement of the first two items by the experimenter is done in the subject's presence, (*c*) whether the experimenter describes to the subject the relation between the first two items, (*d*) whether there is a question at the end of the problem, and (*e*) whether there are insoluble problems. Because of the observed differences in results between Huttenlocher et al., Clark, and the study above due to such procedural variations, and because of the differing conclusions drawn from these results by Clark and by us, it is clear that a decision must be made about what constitutes a fair test of the hypothesized parallel. The fairest test seems to us to be one where the procedure of the arrangement tasks is as identical as possible to that of the syllogisms, as in Huttenlocher et al., where an impressive parallel was found. In any case, since subject-versus-object effects are so sensitive to these various factors in negative-equative arrangement tasks, negative-equatives hardly seem a good sentence form to use in evaluating the existence of a parallel between syllogisms and arrangement tasks.

On Congruence

Clark (1969b) proposes a "congruence" principle to account for certain results in two- and three-term series problems. Recently, he broadened this principle to postulate that in order for a person to answer a question about a sentence, or to verify the truth of a sentence with respect to a state of affairs, the sentence must be "congruent" with the question or the state of affairs in their underlying representations. He has proposed the following four-stage model: Stage 1 involves representation of the "underlying structure" of the sentence. Stage 2 involves representation of the question or state of affairs. Stage 3 involves "comparison operations" where subjects "make underlying representations (i.e., from Stages 1 and 2) completely congruent" so they can make an "identity match" between them. Identity matching of the two inputs is postulated to be essential to the completion of question-answering or verification tasks. Stage 4 involves the choice of an appropriate response (Chase & Clark, 1972; Clark, 1972b).

Clark (1972b) regards question-answering and verification tasks as similar except that, "Whereas Stage 3 simply matched the two representations for identity in the verification task, it must match representations and pick out an X[3] in the question-answering task. It will become clear that this is a minor difference, for the matching operations are carried out in the same way in both tasks." He claims that to evaluate his congruence model, we must consider verification as well as question-answering tasks. As Clark says, his model makes "important empirical claims," first, by asserting how people represent information from sentences and, second, by postulating the existence of a matching stage. We consider these two claims in order.

Representation of information from sentences. The claim that what people store from sentences is their "underlying structures" is a psychological claim which, unfortunately, changes as linguistic theory changes. Clark's (1972b) version holds that the "underlying structure" which people store includes a specification of the lexical items used in a sentence. Thus he admits that the very same facts about the comparative speed of jet planes and rockets may be encoded in the two different comparatives "Jet planes are slower than rockets" and "Rockets are faster than jet planes." Yet he claims that the information from these sentences is stored differently because the sentences involve different adjectives. For instance, he makes the startling claim that, in the former sentence, jets and rockets are both being compared to very fast things and in the latter sentence they are being compared to very slow things. He ignores the importance of topic and context in determing the choice of adjective[4] (see Huttenlocher & Higgins, 1971).

---

[3] Clark's example is the following: "If John killed Harry, then who died?" Stage 1: John caused (Harry die). Stage 2: (X die). Stage 3: Compare (John caused) (Harry die) and (X die: find X).

[4] We do not want to imply that Clark is unaware that topic may sometimes play a significant role in comprehension. In fact, he explicitly discusses topic with respect to comparatives in Clark (1970). However, in his theory of the comprehension of comparatives, he seems to ignore the possibility of a determining role for topic when he says, for instance, that the comparative "A is shorter than B" necessarily implies that both A and B are at the short end of the dimension. Such a position implies that topic could not have dictated the choice of "shorter" as opposed to "taller." That is, one would not, according to Clark, say "A is shorter than B" simply because A was the topic of the immediate conversation even when either A or B or both are *not* at the short end of the dimension. We would argue, on the contrary, that at least for adjectives from a ratio scale, the choice of "shorter" may be dictated by the topic in com-

TABLE 2

PRINCIPLE OF CONGRUENCE AND TEMPORARY ACCESSIBILITY PREDICTED AND ACTUAL DIFFERENCES BETWEEN "BEST" AND "WORST" QUESTION FORMS

Problem	Best–worst predictions		Latency data[a] Clark (1969b)	Success of predictions for (1969b) data		Error data[b] Clark (1969a)	Success of predictions for (1969a) data	
	PC	TA		PC	TA		PC	TA
I (a) $A > B, B > C$	−	−	−.68	*	*	−14	*	*
(b) $B > C, A > B$	−	−	−.54	*	*	−11	*	*
II (a) $C < B, B < A$	+	+	−.26	X	X	−5	*	*
(b) $B < A, C < B$	+	+	+.89	*	*	0	*	*
III (a) $A > B, C < B$	0	+	+.01	*	*	−2	*	*
(b) $C < B, A > B$	0	−	−1.00	X	*	−2	*	X
IV (a) $B < A, B > C$	0	−	−1.02	X	*	+5	X	X
(b) $B > C, B < A$	0	+	+.67	X	*	−11	X	X
I′ (a) $A \not< B, B \not< C$	+	+	+1.82	*	*	+7	*	*
(b) $B \not< C, A \not< B$	+	+	+.60	*	*	+3	*	*
II′ (a) $C \not> B, B \not> A$	−	−	−1.05	*	*	−18	*	*
(b) $B \not> A, C \not> B$	−	−	−.49	*	*	−5	X	X
III′ (a) $A \not< B, C \not> B$	0	−	−.32	*	*	−11	X	*
(b) $C \not> B, A \not< B$	0	+	+.39	X	*	−11	X	X
IV′ (a) $B \not> A, B \not< C$	0	+	−.08	*	*	−4	*	*
(b) $B \not< C, B \not> A$	0	−	−1.64	X	*	−12	X	*

Note.—$>$ = better; $<$ = worse; $\not>$ = not as good; $\not<$ = not as bad; * = correct prediction; X = incorrect prediction; PC = principle of congruity; TA = temporary accessibility. Predictions only diverge for Problems III, IV, III′, and IV′; PC predicts no difference because either both questions are congruent (III, IV′) or incongruent (IV, III′), whereas TA always predicts the problem to be easiest when the adjective in the question matches that in the second premise.

"Worst" questions are harder overall than "best" questions. All comparisons are made from corrected scores obtained by averaging across all problem types; corrected reaction time = −.17 seconds, corrected error = −5.7%. Directionally correct predictions given *. For 0 differences predictions, * given when reaction times = −.17 ± .15 and errors = −5.7% ±5%. These 16 scores of differences between "best" and "worst" question forms provide no basis for estimating confidence limits for Clark's zero predictions. The reason one cannot obtain an estimate of variance of range for zero predictions from these differences is that, by *our* theory, they all represent genuine population differences of varying sizes. Even by Clark's theory, half of these scores should represent genuine population differences. To estimate variation for his zero predictions by using the eight scores where Clark's theory predicts no differences would, of course, be absurd, since it presupposes that these are indeed zero differences, thus assuming a priori what one is trying to test. To estimate variation for zero predictions, one needs within-group scores. Since only between-group data are available from Clark's publications, we examined our own data for several groups doing three-term series problems. We based these confidence intervals, ±.15 and ±5%, on differences we have found to be significant at .01 approximately half of the time using Mann-Whitney U's on untransformed data. While these confidence limits are generous to PC, the wary reader can instead examine the sizes of observed differences for these zero predictions relative to the entire set of 16, half of which, by Clark's own theory, are genuine differences. Some of the largest observed differences fall in categories where he predicts zero differences. While an avid supporter of PC might not assign an X to reaction time differences for IVb or III′b, or to error differences for IVa, he surely would assign an X to reaction time differences for IIIb, IVa, and IV′b, and to error differences for IVb, III′a, III′b, and IV′b.

[a] Differences for geometric mean times in seconds.

[b] Differences for errors in percentages.

Taking Clark's (1972b) model literally, the information about the world which people obtain from sentences remains forever framed in terms of the "underlying structures" of those sentences, including the particular lexical items used. Even if Clark does favor the notion that information in long-term memory is represented in some linguistic form, surely he would want to allow it some other organization than a list of the "underlying structures" of input sentences. Even the weaker claim that information from a sentence is represented in terms of its "underlying structure" for some limited time period before it becomes part of a person's general store of information about events must be evaluated against alternative possibilities.

Clark (1969b) did not evaluate his congruence principle against even the most obvious alternatives. His evidence was from two-term and three-term series problems. Two-term series problems, for example, "Sam is better (worse) than Joe. Who is best (worst)?" are easiest when the adjective in the question matches that in the sentence. This finding, however, can be explained without making any "congruence" claims about information storage. One alternative is that "match" cases are easier because, for such problems, subjects do not obtain a semantic interpretation of the sentence at all, but only apply certain syntactic knowledge, as they might for a problem like "Sam is glupper than Joe. Who is gluppest?" A sec-

---

bination with a knowledge of the relative height of A and B on the scale, and *for this very reason* does not imply that A and B are at the short end of the dimension.

ond alternative is that, for "match" cases, memory-search time is minimized. That is, one set of operations which may take time in obtaining semantic interpretations of sentences is searching for the meanings of the words. If the subject has just "looked up" a word, its meaning might temporarily become more accessible.

We used Clark's three-term series data to test his principle of congruence against just such a "temporary accessibility" alternative, which said that what was important was whether the adjective in the question matched the adjective in the second premise, not whether it matched the adjective used to present information about an item in the first place. His 1969b data were consistent with temporary accessibility, *not* with the principle of congruence: When the answer was in the first premise, a match between the adjective in the question and the adjective in that first premise did not help but, rather, questions were easier when the adjective in the question matched that in the *second* premise.

Clark's rebuttal claims that (*a*) his 1969b data, the supposed basis for his theory, are based on only "13 subjects with two observations per subject per problem" and (*b*) his 1969a data are "more precise." We had posed temporary accessibility as a counterhypothesis against which to evaluate "congruence." Since *neither* temporary accessibility nor the principle of congruence is very good at handling the 1969a data, we retain our original conclusion that there is no strong evidence for "congruence." First, contrary to both the principle of congruence and temporary accessibility, Clark's 1969a data show that when the answer is in the first premise, and the same adjective appears in both premises, a problem is easier when the adjective in the question does *not* match that in the premises (24.7% errors) than when it does match (31.4%). Second, a systematic comparison of the principle of congruence and temporary accessibility for all soluble problems from both 1969a and 1969b, as shown in Table 2, indicates that neither is a powerful explanatory principle.

Matching. With respect to verification tasks, Clark (1972b) claims that "It is obvious that there must be a representational system common to perception and language; otherwise, it would be impossible for people to talk about pictures" and that "it is necessary to assume that pictures are represented at some level of processing in semantic representations like those of sentences." This matching stage was postulated to answer the question: How does a listener determine whether a sentence accurately describes a state of affairs? It answers the question as follows: The listener makes his decision by formulating a sentence about it (or a semantic representation like a sentence) and by determining if that sentence matches the input sentence (or its semantic representation). Contrary to Clark, we believe that it is unlikely that such a matching stage exists at all except, perhaps, as a special strategy under unusual conditions.

To make our point, we consider the simplest type of verification task—where a person must decide whether a word is appropriately applied to a picture; for example, the word "cow" to a picture of a cow. Figure 1 shows what is required to make such a decision. The subject must be able to recognize a visual display of a cow as a member of a particular perceptual

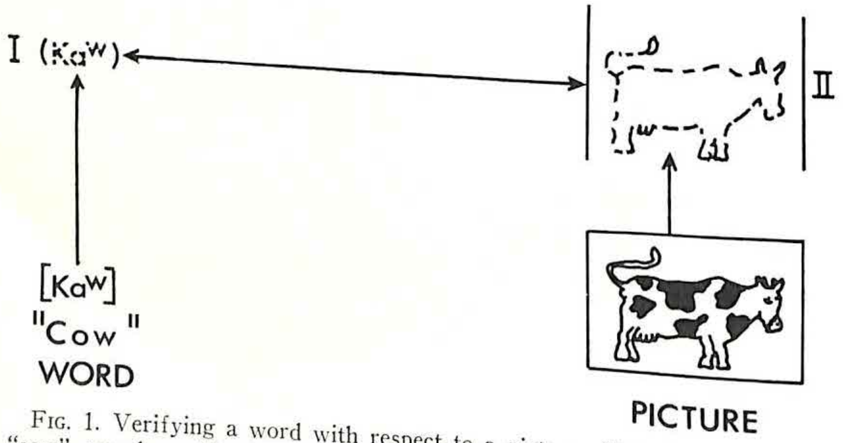


FIG. 1. Verifying a word with respect to a picture. (I = prototype of the "cow" sound combination; II = prototype of the "cow" visual representation.)

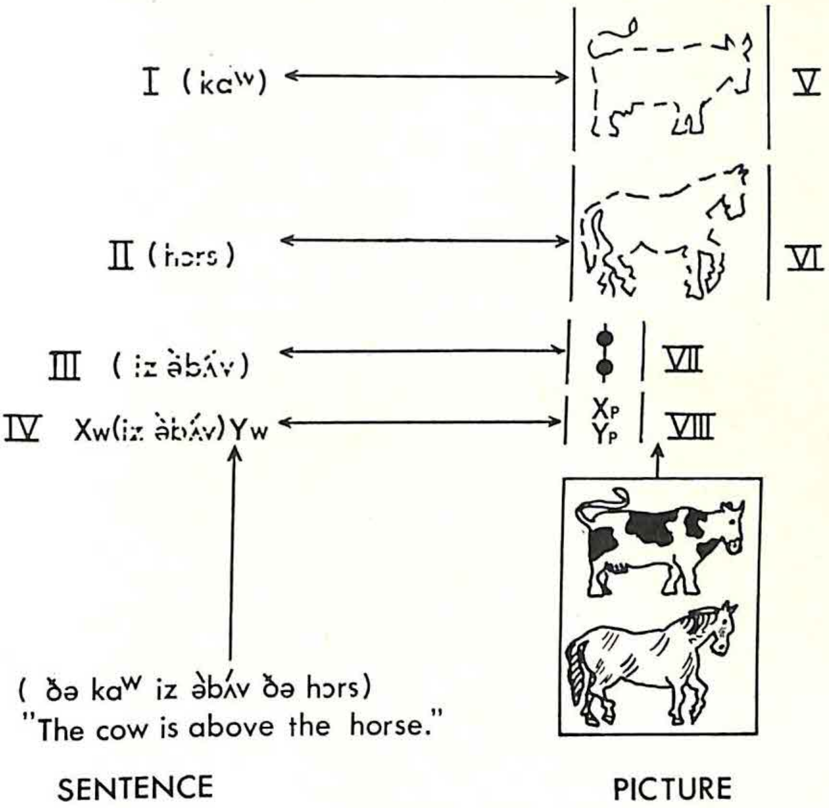


FIG. 2. Verifying a sentence with respect to a picture. (I = prototype of the "cow" sound combination; II = prototype of the "horse" sound combination; III = prototype of the "is above" sound combination; IV = prototype of the "X_w is above Y_w" syntactic relation where X_w and Y_w are both words with privilege of occurrence before and after "is above"; V = prototype of the *cow* visual representation; VI = prototype of the *horse* visual representation; VII = prototype of the *vertical relation* visual representation; VIII = prototype of the *vertical order with X_p on top and Y_p underneath* visual representation where X_p and Y_p are both perceptual entities.)

category (perceptual prototype). Similarly, the word "cow" must be recognized as a set of sounds which forms the sound pattern of a particular English word (word prototype). In addition, and most important, the subject must have stored a mapping of the perceptual prototype of the cow onto the word prototype "cow." To decide if the word can be appropriately applied to the picture, the subject must categorize each input into its appropriate prototype and must determine whether or not the appropriate mapping exists.[5] The same features are present when the subject decides whether the sentence "The cow is above the horse" is appropriately applied to a picture of a cow above a horse. Figure 2 shows what is required to make such a decision. The subject can test whether each word is used appropriately, as in the "cow" example above. In addition, he must recognize the display as belonging to the category of a particular type of perceptual relation between items (perceptual-relation prototype) and as involving a particular order of those items. The relation among words in the sentence must also be categorized (syntactic prototype). To decide if the sentence is appropriate to the picture, subjects must determine if there exists a mapping of this ordered relation among objects onto this ordered relation among words.[6]

[5] The notion of a perceptual prototype is derived from the work of Posner (1969) and a fuller discussion of our use of this notion in respect to an analysis of symbols will appear in E. T. Higgins and J. Huttenlocher, Symbols and other signs: Their characteristics and functions in mental processes. In preparation.

[6] We obviously cannot do justice to the complicated matter of the relation between surface and

Clark's (1972b) model postulates instead that subjects must match an input word or sentence against an internally generated word or sentence about the picture, rather than applying directly a knowledge of how the lexicon and syntactic rules of the language are mapped onto perceived objects and relations among objects. That is, the subject only knows that the word "cow" applies to a cow picture by internally generating the word "cow" from the picture, and then matching it to the input word. Surely though, such internal generation cannot occur unless the subject knows that such a word maps onto such a picture. Thus it is utterly mysterious why the subject cannot simply check whether the appropriate mapping exists, rather than having to take the extra steps of internal generation and identity matching. According to such a theory, if a person could not *recall* an object's name then he could not *recognize* that name as appropriate to the object even if someone told it to him, because judging the applicability of a word to an object is held to depend on having internally generated the object's name, that is, on having recalled it. But there is considerable empirical evidence that people can recognize even when they cannot recall, as, for example, in the tip-of-the-tongue phenomenon.

With respect to the verification of sentences, Chase and Clark (1972) say of "picture first" tasks that when the picture is naturally encoded as "A above B," and the input sentence is "B below A," then even though the sentence is true of the picture, "This cannot be discovered directly by a match–mismatch comparison. People therefore, first make sure that the subjects of the picture and sentence representations are identical. . . ."[7] In order to change "A is above B" to "B is below A," subjects must know that sentences with "above" are synonymous with (map onto) sentences with "below" if one reverses the order of A and B. Again, it is mysterious why, given that the semantic representation of one statement involves "above" and that of the other statement involves "below," the subject cannot simply check whether there exists the appropriate reversal of item order (i.e., whether the order of A and B is reversed).

For question answering, the model is equally gratuitous for exactly the same reason. It is claimed that sentence and question must be identical to solve for X in the question (see Footnote 3). We would argue instead that rather than it being necessary to match the underlying representations of question and sentence as part of the solution process, it is simply necessary to know that expressions involving the word "kill," which describe particular logical relations among items, map onto other expressions involving the word "dead" which describe those same logical relations.

In sum, whether or not one postulates *matching* operations, one *must* postulate mapping operations. That is, subjects can only talk about objects and events if they know how the symbols and the relations among symbols in the language are mapped onto perceived objects and relations among objects. Furthermore, subjects can only turn one expression into another expression if they know how different linguistic forms are mapped onto one another. It is more economical to simply postulate that subjects check whether such rules of relation exist in particular cases than that they turn one input into the same form as that of the other and then match the two. Therefore, one wants strong empirical evidence before postulating that such a matching stage exists at all. There is no such evidence at present.

Furthermore, it is difficult to understand how one even *could* build a satisfactory model which postulates such identity matching except for tasks which are artificially constrained. After all, the mapping of language onto the world of objects and events is not one-to-one. That is, a particular sentence can be applied to many perceptual situations, and a particular perceptual situation may be described by many different sentences (sentences with different "underlying structures"). The problem is especially obvious in "picture first" tasks. For example, a star and circle in vertical array might be encoded "pointed object is near round object" as well as "star above circle" if subjects were not limited to eight possible sentences. Finally, one must question the whole notion that the information from pictures is always encoded in "semantic representations like those of sentences."

On Lexical Marking

Clark (1969) proposed that comparatives are easier to understand with unmarked than

---

underlying syntax and their mappings onto the perceptual relations among objects within our present space restrictions. A further discussion of what is involved will appear in J. Huttenlocher and E. T. Higgins, Language and other symbolic schemes: Their use in representing information. In preparation.

[7] Regarding "the subjects," Clark refers here to sentence subjects, that is, the subject of an "underlying representation."

with marked adjectives because only the former have a neutral "nominal" sense, and that "an unmarked adjective used in its 'nominal' sense is easier to comprehend than an unmarked or marked adjective used in its 'contrastive' sense [p. 389]." Thus he claims that unmarked adjectives can be used neutrally in comparatives, whereas marked adjectives cannot. Huttenlocher and Higgins (1971) pointed out that this proposal (*a*) is inconsistent with speakers' intuitions that comparatives like "D is taller than E" and "E is shorter than D" are equally neutral, (*b*) is not parsimonious because it cannot explain observed differences in the difficulty of simple contrastive use of marked and unmarked adjectives, and (*c*) has no independent empirical support.

Clark (1971) claims that his hypothesis *does* have empirical support from a study by Jones (1970) which contrasted syllogisms with polar pairs which either did or did not contain an unmarked member, that is, heavy–light and thick–thin versus dark–light and fat–thin. In comparing syllogisms containing the same adjective in both premises, Jones found that for the first two but not the last two polar pairs, the member adjectives differed significantly in difficulty. This is not very impressive evidence. First, to show that unmarked–marked pairs differ more in difficulty than other polar pairs, one must determine the mean and variability of differences on the basis of *many* such unmarked–marked pairs, not just cite two cases. We did Jones' comparison for the two unequivocally unmarked–marked pairs from Huttenlocher et al. (good–bad, high–low). For the latter pair, the member adjectives do *not* differ from one another significantly in difficulty. Second, the use of polar pairs where the marked adjective is the "same" word (light, thin), which is contrasted with two different types of polar opposites, seems to suggest that any variation in the size of differences would be due to "heavy" and "thick," the unmarked adjectives, being easier than "dark" and "fat," the marked adjectives. However, for "thin" at least, Jones' results show that the effect instead derives chiefly from the differences in the difficulty of the two senses of "thin."

Finally, Clark (1971) criticizes our use of the terms "unmarked" and "marked." Anyone who has examined the linguistic literature of the last decade knows that these terms do not, as Clark claims, have any universally agreed-upon definition. This necessarily leads to some confusion about whether or not particular polar pairs should be regarded as unmarked–marked pairs. Clark and Card (1969), in a study with eight polar pairs, used three questionable pairs: hot–cold, dirty–clean (neither pair, Clark now correctly argues in criticizing us for using them, has an unmarked member), and pretty–ugly. Later in that same paper, Clark and Card suggest "warm" be substituted for "hot" and "beautiful" for "pretty." While the former is an improvement, "beautiful' is surely no more neutral than "pretty." Thus, at least in 1969, Clark had difficulty determining what constitutes an unmarked–marked pair.

REFERENCES

CHASE, W. G., & CLARK, H. H. Mental operations in the comparison of sentence and pictures. In L. Gregg (Ed.), *Cognition in learning and memory.* New York: Wiley, 1972, in press.

CLARK, H. H. The influence of language in solving three-term series problems. *Journal of Experimental Psychology*, 1969, **82**, 205–215. (a)

CLARK, H. H. Linguistic processes in deductive reasoning. *Psychological Review*, 1969, **76**, 387–404. (b)

CLARK, H. H. Comprehending comparatives. In G. B. F. D'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics.* New York: American Elsevier, 1970.

CLARK, H. H. More about "Adjectives, comparatives, and syllogisms": A reply to Huttenlocher and Higgins. *Psychological Review.* 1971, **78**, 505–514.

CLARK, H. H. Difficulties people have answering the question "Where is it?" *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 265–277. (a)

CLARK, H. H. Semantics and comprehension. In T. A. Sebeok (Ed.), *Current trends in linguistics.* Vol. 12. *Linguistics and adjacent arts and sciences.* The Hague: Mouton, 1972, in press. (b)

CLARK, H. R., & CARD, S. K. Role of semantics in rememberering comparative sentences. *Journal of Experimental Psychology*, 1969, **82**, 545–553.

HUTTENLOCHER, J. Constructing spatial images: A strategy in reasoning. *Psychological Review*, 1968, **75**, 550–560.

HUTTENLOCHER, J., & HIGGINS, E. T. Adjectives, comparatives, syllogisms. *Psychological Review*, 1971, **78**, 487–504.

HUTTENLOCHER, J., HIGGINS, E. T., MILLIGAN, C., & KAUFFMAN, B. The mystery of the "negative equative" construction. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 334–341.

JONES, S. Visual and verbal processes in problem-solving. *Cognitive Psychology*, 1970, **1**, 201–214.

POSNER, M. I. Abstraction and the process of recognition. In J. T. Spence & G. H. Bower (Eds.), *Psychology of learning and motivation.* Vol. 3. New York: Academic Press, 1969.

(Received November 22, 1971; revision received March 24, 1972)

Psychological Review
1972, Vol. 79, No. 5, 428–432

ON THE EVIDENCE CONCERNING J. HUTTENLOCHER AND E. T. HIGGINS' THEORY OF REASONING:

A SECOND REPLY[1]

HERBERT H. CLARK[2]

Stanford University

In answering H. H. Clark's reply to their earlier paper, J. Huttenlocher and E. T. Higgins dispute the counterevidence to their imagery theory of reasoning and offer arguments against Clark's previous linguistic theory of reasoning. In the present reply, it is argued that the counterevidence to the imagery theory remains valid since (*a*) the only evidence supporting the imagery theory is suspect because of an experimental confounding and (*b*) there are several independent findings contradictory to the imagery theory. It is also contended that the arguments against the linguistic theory are unfounded. Finally, new evidence is presented against Huttenlocher and Higgins' explanation for the advantage of "unmarked" over "marked" comparatives, for example, *longer* over *shorter*.

In this reply, I will consider Huttenlocher and Higgins' (1972) major counterarguments to my earlier reply (Clark, 1971) to their previous proposals (Huttenlocher & Higgins, 1971) about how people solve three-term series problems (e.g., "If John is better than Pete, and John is worse than Dick, then who is best?"). I will concentrate on the pertinent empirical evidence.

Counterevidence to the Imagery Theory

The imagery theory of reasoning (Huttenlocher, 1968; Huttenlocher & Higgins, 1971) claims, roughly, that people solve three-term series problems by manipulating mental objects in mental arrays. Thus, it predicts a parallel between the difficulties people have in placing real objects into real arrays (placement tasks) and the difficulties people have in solving reasoning problems (reasoning tasks). In a direct test of this theory, Clark (1972a) found strong evidence against this parallel. From their own slightly different tests, however, Huttenlocher, Higgins, Milligan, and Kauffman (1970) and Huttenlocher and Higgins (1972) concluded that there *was* a parallel. Which of these two conflicting tests is correct?

A critical difference between the two tests is in the type of placement task used. Clark (1972a) used a *pure placement task*. In it the subject read an instruction like "C is lower than B" (corresponding to the second premise in a reasoning task). He then viewed an array already containing, say, A and B and placed the third object (C) into the array according to the instruction. Huttenlocher and her colleagues used a *mixed placement task*, consisting of two stages. At the first stage, the subject heard an instruction like "A is higher than B" (corresponding to the first premise in a reasoning task) and then placed A and B into an otherwise empty array. The second stage was an equivalent of the pure placement task. The timing in both the pure and mixed placement tasks covered only the final placement of the third object (here, C). Despite these similarities, the mixed placement results were generally consistent with the reasoning results, whereas the pure placement tasks were not.

As I have argued previously (Clark, 1971, 1972a), the mixed placement task contains a probable confounding, making it invalid as a test of the imagery theory. When the subject is given both the first and second premises of a reasoning problem, as in the mixed placement task, he can use either one of two strategies: He could ignore the first premise altogether and deal with the second premise exactly as he would in a pure placement task (Strategy 1), or he could make use of the first and second premises together, "reason" with them to discover the final ordering of the three objects, and then place the third object in accordance with this

---

[1] The preparation of this reply was supported in part by United States Public Health Service Research Grant MH-20021 from the National Institute of Mental Health. I wish to thank P. Carpenter, E. V. Clark, M. Just, R. Revlis, and E. E. Smith for their helpful comments on earlier drafts of this paper.

[2] Requests for reprints should be sent to Herbert H. Clark, Department of Psychology, Stanford University, Stanford, California 94305.

reasoning (Stategy 2).[3] The mixed placement results, therefore, should be identical to the pure placement results for those subjects using Strategy 1, but to the reasoning results for those subjects using Strategy 2. Indeed, most (though not all) of the mixed placement results were parallel to the reasoning results, suggesting that most subjects had used Strategy 2. But when Strategy 2 is used, the mixed placement task is no longer really a "placement" task, but is equivalent to the reasoning task it is trying to predict. In contrast, the pure placement task, which meets all other criteria for testing the imagery theory,[4] contains no such confounding, and its results go counter to the theory.

A further difficulty for Huttenlocher and Higgins is that they are left with no explanation for the discrepant results of the two placement tasks. The linguistic theory in Clark (1969b), however, can account for this disparity. It correctly predicts one particular pattern for the pure placement tasks (see Clark, 1972a), but quite a different pattern for pure reasoning problems (see Clark, 1969b). Insofar as subjects in the mixed placement tasks are reasoning (using Strategy 2), the mixed placement results should parallel the reasoning results, not the pure placement results. Indeed, this is approximately what happens. In summary, the evidence against the imagery theory numbered 1 in Clark (1971) still stands.

Even if the placement tasks of Huttenlocher et al. (1970) could be accepted as a proper test of the imagery theory, they do not show the "impressive parallel" with reasoning tasks that Huttenlocher and Higgins claim. Huttenlocher et al., for example, compared placement and reasoning problems whose premises contained *high* and *low*. If the imagery theory were correct, the difference scores summarized in Table 1 should be parallel in the two tasks (compare columns 3 and 5). But as the scores show, the latency and error differences were large and significant in the placement task, but zero or negligibly small in all but one instance in the reasoning task.[5] These results could even be construed as disconfirming the imagery theory.

TABLE 1

LATENCY AND ERROR DIFFERENCES IN HUTTENLOCHER ET AL.'S (1971) HIGH–LOW PLACEMENT AND REASONING TASKS

Measure of difficulty used	Type of problem	Type of task			
		Placement	*p*	Reasoning	*p*
Latencies[a]	Comparative	670	.01	20	*ns*
	Negative-equative	−240	.07	−20	*ns*
Percent errors	Comparative	12	.001	0	*ns*
	Negative-equative	12	.001	5	.03

Note.—Each difference was calculated by subtracting the mean (latency or percent error) for problems in which the third object to be placed is the subject of the second premise from the mean for problems in which the third object is in the predicate of the second premise. The significance level for each difference is given in columns 4 and 6.

[a] In milliseconds.

In the first reply (Clark, 1971, p. 507) and in Clark (1972a), I discussed a second independent type of counterevidence to the imagery theory. According to the imagery theory, the so-called "indeterminate" reasoning problems of Clark (1969a) should result in the same pattern of difficulties as the "determinate" reasoning problems. Since they clearly did not, the imagery theory must be incorrect. In their reply, Huttenlocher and Higgins (1972) do not take up this counterevidence,[6] hence the conclusion still stands.

Principle of Congruence

In reconsidering the principle of congruence (that a statement and question asked of it must match in their representations) and their own alternative notion of "temporary accessibility" (that the meaning of a word becomes more available for a short time after its use),

---

[3] This process is described more completely in H. H. Clark, More on linguistic processes in deductive reasoning. In preparation.

[4] In her original demonstrations of the imagery theory, Huttenlocher (1968) relied on pure, not mixed, placement tasks.

[5] Actually, since latencies and errors go in opposite directions on the negative-equative problems of the placement task in Table 1, it seems indefensible to draw any conclusions from the data (see Yellott, 1971). The same difficulty reappears in Huttenlocher and Higgins' (1972) placement task. Their Table 1 shows significant differences in latencies on eight comparisons; yet on five of the eight, the error differences are from 3% to 12% ($M = 7.4\%$) in the opposite direction. The errors in Clark's (1972a) tests were much rarer and did not contradict the latencies.

[6] Huttenlocher and Higgins (1972) say that negative-equative constructions "hardly seem a good sentence form to use" in evaluating the imagery theory. This counterevidence, however, does not rely on negative-equative constructions.

Huttenlocher and Higgins (1972) concur with the first reply (Clark, 1971) that the available evidence on reasoning goes counter to "temporary accessibility," but they now claim there is evidence against congruence too. This claim, however, is based on two illegitimate tests. First, Huttenlocher and Higgins assert (see their Table 2) that congruence predicts latency and error differences of zero between "Who is best?" and "Who is worst?" on Clark's (1969a, 1969b) problems III, IV, III′, and IV′. This is incorrect. In these individual instances, congruence makes no predictions, and this does not necessarily imply there should be zero differences on each problem (other factors could be involved). Without the illegitimate comparisons, their Table 2 supports congruence on 14 of 16 individual predictions, with the two remaining differences insignificantly small. Second, Huttenlocher and Higgins make note of a particular 24.7% to 31.2% error difference supposedly contradictory to congruence. Clark (1969a, 1969b) and Huttenlocher (1968), however, have both found that when there is the same adjective in both premises, syllogisms are consistently easier with their two premises in one particular order than in the other. Huttenlocher and Higgins' comparison is confounded with this independent factor. All comparisons not confounded by such factors are consistent with congruence.

On the basis of these tests alone, Huttenlocher and Higgins (1972) say, "We retain our original conclusion that there is no strong evidence for 'congruence' [p. 424]." They have drawn this conclusion, however, without addressing themselves to all the other evidence that has been claimed to support congruence or matching, for example, by Chase and Clark (1972), Clark (1972a, 1972b), Clark and Chase (1972), Ellis (1972), Just and Carpenter (1971), Marks (1972), Olson and Filby (1972), Seymour (1969, 1971), Trabasso (1972), and Trabasso, Rollins, and Shaughnessy (1971).

One of Huttenlocher and Higgins' (1972) main criticisms of the Chase and Clark (1972) theory of sentence-picture comparison (hence the principle of congruence) is illustrated in the following comment: "According to such a theory, if a person could not recall an object's name then he could not recognize that name as appropriate to the object even if someone told it to him, because judging the applicability of a word to an object is held to depend on having *internally generated the object's name* [italics added, p. 426]." But as explicitly stated in the Chase and Clark proposal, the subject does not necessarily generate an internal name for an object; rather, he constructs a "semantic representation" for it, a code Chase and Clark carefully distinguish from internally generated names. Once this misconception is corrected, this and other objections by Huttenlocher and Higgins no longer apply.

As an alternative to matching or congruence models, Huttenlocher and Higgins outline their own mapping model. In it the subject is viewed as trying to find a "mapping" from sentences onto situations they describe. Alone, however, the model makes no predictions and cannot be evaluated. Presumably, therefore, it is meant to subsume a "correspondence" model that Huttenlocher and her colleagues had previously proposed to account for placement and reasoning tasks (Huttenlocher, 1968, pp. 553–554; Huttenlocher, Eisenberg, & Strauss, 1968, p. 528; Huttenlocher & Strauss, 1968, p. 304; Huttenlocher & Weiner, 1971, p. 373). If so, however, serious contradictions arise.

First, in criticizing the congruence models, Huttenlocher and Higgins (1972) argue that it is not necessary for the subject to change one representation into another to be able to compare a sentence and the situation it describes. Yet the "correspondence" model does just that. As Huttenlocher (1968) described the model, "Comprehension requires a correspondence between the form of a linguistic expression and the situation it describes [p. 553]"; when there is no correspondence, the subject must "carry out mental operations to create a correspondence" between the two by, for example, "transforming" one or the other representation (see Huttenlocher, 1968, p. 554).[7] Indeed, it is these mental operations that enable one to predict which placement and reasoning problems will be hard and which will not. If the mapping model excludes such mental operations, as it appears to, then it cannot make those predictions.

Second, regardless of the technical differences between the "correspondence" and "congruence" models, each instance for which Huttenlocher and her colleagues have claimed a lack of correspondence is also an instance of lack of congruence (see Clark, 1972a, 1972b).

---

[7] Compare this description of "correspondence" to Clark's (1972b) description of "congruence": "The principle of congruence asserts that two such semantic representations [as of a sentence and the situation it describes] are compared for congruence, and in the absence of congruence, there must be mental operations that bring the two representations into congruence."

As a result, every difference predicted by correspondence would also be predicted by congruence. Huttenlocher and Higgins (1972), therefore, cannot conclude "that there is no strong evidence for 'congruence' [p. 424]" without contradicting their own previous evidence for correspondence.[8]

Principle of Functional Relations

Huttenlocher and Higgins (1971, 1972) have claimed that antonymous comparatives on ratio scales (e.g., *higher* and *lower*) do not differ as to where they can be applied on their underlying scales (e.g., height) (see Clark, 1971, pp. 510–512). Contary to this claim, however, Audley and Wallis (1964) have found that it is easier to answer "Which is brighter?" than "Which is darker?" for two very bright lights differing slightly in brightness, but just the reverse for two very dark lights differing slightly in brightness. The principle of congruence predicts such a congruity effect. According to Huttenlocher and Higgins' hypothesis, however, the two questions should be parallel in difficulty regardless of absolute brightness, and this is contrary to fact. The same disconfirmatory result has been found for *higher–lower* (Wallis & Audley, 1964), *older–younger* (Ellis, 1972), and *more–less* (Marks, 1972).

Principle of Lexical Marking

Huttenlocher and Higgins (1972) also reconsider two hypotheses as to why unmarked comparatives like *better* and *longer* are easier than marked comparatives like *worse* and *shorter*. Briefly, the neutralization hypothesis (Clark, 1969b) states that *better* is easier than *worse* because *better* neutralizes and *worse* does not. The negativity hypothesis (Clark, 1970, 1972b; Huttenlocher & Higgins, 1971) states that *worse* is more difficult than *better* because *bad* is implicitly negative and *good* is not. In Clark (1971, 1972b), I argued that both hypotheses are still viable, but Huttenlocher and Higgins continue to claim that the negativity hypothesis is the only one that is correct. A recent study (Clark, unpublished), however, appears to disconfirm their claim. In that study, subjects were timed as they made true–false judgments of sentences like "The blue line is long/short" and "The blue line is longer/shorter than the pink line" against appropriate pictures of blue and pink lines. The negativity hypothesis predicts a latency advantage of *long* over *short* and, because no other factors are involved, the identical latency advantage of *longer* over *shorter*. The neutralization hypothesis, however, predicts an advantage of *longer* over *shorter* over and above the advantage of *long* over *short*. This is because *longer* is neutral as well as not being negative. The results showed a 42-millisecond ($p < .05$) advantage of *long* over *short*, but a significantly larger ($p < .01$) 116-millisecond advantage of *longer* over *shorter*. These findings are consistent with the two hypotheses independently of each other and are inconsistent with Huttenlocher and Higgins' claim that the latency advantage of *longer* over *shorter* can be accounted for solely by the negativity hypothesis.

Conclusion

In the limited space available for this reply, I have argued that the previous counterevidence to the imagery theory (see Clark, 1969b, 1971, 1972a, 1972b) remains valid despite Huttenlocher and Higgins' (1972) counterarguments. Furthermore, the linguistic theory still offers a viable account for the solution of three-term series problems (see Clark, 1969a, 1969b, 1970, 1971, 1972b). It has also been extended to account for placement tasks (Clark, 1972a, 1972b). It should be reemphasized, however, that the evidence presented here does not rule out all imagery theories of reasoning: it speaks only to Huttenlocher and Higgins' particular formulation of such a theory.

REFERENCES

AUDLEY, R. J., & WALLIS, C. P. Response instructions and the speed of relative judgments: I. Some experiments on brightness discrimination. *British Journal of Psychology,* 1964, **55,** 59–73.

CAREY, S. The syntactic and referential aspects of linguistic encoding. Unpublished honor's thesis, Harvard University, 1964.

CHASE, W. G., & CLARK, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), *Cognition in learning and memory.* New York: Wiley, 1972.

CLARK, H. H. The influence of language in solving three-term series problems. *Journal of Experimental Psychology,* 1969, **82,** 205–215. (a)

CLARK, H. H. Linguistic processes in deductive reasoning. *Psychological Review,* 1969, **76,** 387–404. (b)

---

[8] Similarly, in discussing the "correspondence" model, Huttenlocher and Strauss (1968, p. 304) endorse Carey's (1964) "matching" model for the verification of sentences against pictures. Yet Huttenlocher and Higgins (1972) object to Chase and Clark's (1972) virtually identical "matching" model for the same task.

CLARK, H. H. How we understand negation. Paper presented at the Committee on Basic Research in Education workshop on cognitive organization and psychological processes. Huntington Beach, California, August 1970.

CLARK, H. H. More about "Adjectives, comparatives, and syllogisms": A reply to Huttenlocher and Higgins. *Psychological Review,* 1971, **78**, 505–514.

CLARK, H. H. Difficulties people have in answering the question "Where is it?" *Journal of Verbal Learning and Verbal Behavior,* 1972, **11**, 265–277. (a)

CLARK, H. H. Semantics and comprehension. In T. A. Sebeok (Ed.), *Current trends in linguistics.* Vol. 12. *Linguistics and adjacent arts and sciences.* The Hague: Mouton, 1972. (b)

CLARK, H. H., & CHASE, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology,* 1972, in press.

ELLIS, S. H. Interaction of encoding and retrieval in relative age judgments: An extension of the "crossover" effect. *Journal of Experimental Psychology,* 1972, in press.

HUTTENLOCHER, J. Constructing spatial images: A strategy in reasoning. *Psychological Review,* 1968, **75**, 550–560.

HUTTENLOCHER, J., EISENBERG, K., & STRAUSS, S. Comprehension: Relation between perceived actor and logical subject. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7**, 300–304.

HUTTENLOCHER, J., & HIGGINS, E. T. Adjectives, comparatives, and syllogisms. *Psychological Review,* 1971, **78**, 487–504.

HUTTENLOCHER, J., & HIGGINS, E. T. On reasoning, congruence, and other matters. *Psychological Review,* 1972, **79**, 420–427.

HUTTENLOCHER, J., HIGGINS, E. T., MILLIGAN, C., & KAUFFMAN, B. The mystery of the "negative equative" construction. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9**, 334–341.

HUTTENLOCHER, J., & STRAUSS, S. Comprehension and a statement's relation to the situation it describes. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7**, 527–530.

HUTTENLOCHER, J., & WEINER, S. L. Comprehension of instructions in varying contexts. *Cognitive Psychology,* 1971, **2**, 369–385.

JUST, M., & CARPENTER, P. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10**, 244–253.

MARKS, D. F. Relative judgment: A phenomenon and a theory. *Perception and Psychophysics,* 1972, **11**, 156–160.

OLSON, D. R., & FILBY, N. On the comprehension of active and passive sentences. *Cognitive Psychology,* 1972, in press.

SEYMOUR, P. H. K. Response latencies in classification of word-shape pairs. *British Journal of Psychology,* 1969, **60**, 312–321.

SEYMOUR, P. H. K. Perceptual and judgmental bias in classification of word-shape displays. *Acta Psychologica,* 1971, **35**, 461–477.

TRABASSO, T. Mental operations in language comprehension. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* Washington, D.C., Winston, 1972.

TRABASSO, T., ROLLINS, H., & SHAUGHNESSY, E. Storage and verification stages in processing concepts. *Cognitive Psychology,* 1971, **2**, 239–289.

WALLIS, C. R., & AUDLEY, R. J. Response instructions and the speed of relative judgments: II. Pitch discrimination. *British Journal of Psychology,* 1964, **55**, 121–132.

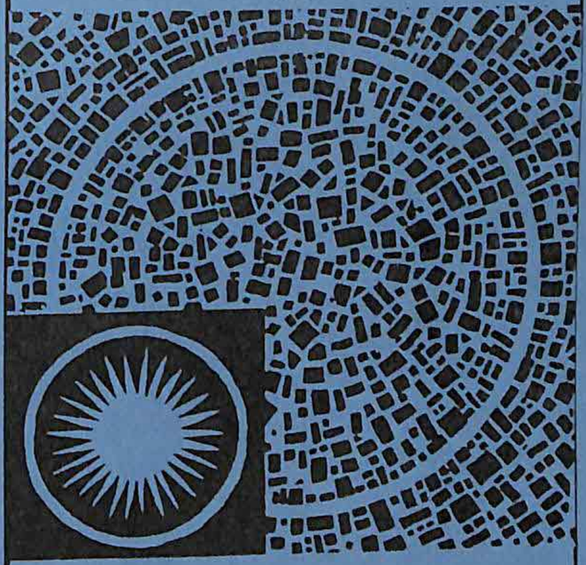
YELLOTT, J. I. Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology,* 1971, **8**, 159–199.

(Received February 22, 1972; revision received May 8, 1972)

The *Psychological Review* is initiating blind reviewing of manuscripts on an experimental basis. Beginning immediately, title pages, abstracts, footnotes, or manuscript pages should not show any identification of the paper's author or authors. With each copy of the manuscript, authors should submit an additional title page which shows the title of the paper, name of the author(s), and institutional affiliation. Manuscripts should be submitted in triplicate as in the past. Proper preparation of manuscripts is the author's responsibility. The system of blind reviewing does *not* apply to Letters to the Editor.

"I think that Masters and Houston actually have broken through to a new understanding of the sense and uses of the disciplines of inward-turned contemplation—an understanding that leaves the Freudian schools of technique and theory far behind."

—JOSEPH CAMPBELL



mind games

by Robert Masters
and Jean Houston

Two eminent mind researchers show how an individual can realize his full potential *without* the use of drugs. How he can examine both the outer limits and inner reaches of his imagination and his physical world, and gain a complete and rewarding sense of mind and body. How he can reach and be the best of himself, safely.

"I have read three important and revolutionary books in the last three years: Yoko Ono's *Grapefruit*, Arthur Janov's *Primal Scream*, and now *Mind Games*. I suggest you read and experience them."

—JOHN LENNON

$7.95

THE VIKING PRESS
625 Madison Avenue
New York, N.Y. 10022

2 FEB 1974

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

announces new developments in

Psychological Abstracts Information Services

The APA has been expanding the *Psychological Abstracts (PA)* information service offered to the psychological community since *PA*'s inception in 1927 and NOW offers the capability for automated search, in a matter of minutes, of all records published in *PA* over the last 5 years. This new development may revolutionize the researcher's use of the psychological literature. Search and retrieval is now as easy as 1, 2, 3 . . .

1 BEGIN SEARCH.

2 REQUEST: *PSYCHOLOGICAL ABSTRACTS* INFORMATION SERVICES◣

3 8 ENTRIES. DO YOU WANT TO PROCESS ANSWERS: NO, PRINT, OR MODIFY?

4 MODIFY – AUTOMATED SEARCH AND RETRIEVAL SERVICES◣

5 3 ENTRIES. DO YOU WANT TO PROCESS ANSWERS: NO, PRINT, OR MODIFY?

6 PRINT◣

7 ENTER DESIRED OUTPUT.

8 FULL SUMMARY◣

9 PASAR (*PA* SEARCH AND RETRIEVAL), A CENTRALIZED AUTOMATED SERVICE THAT PROVIDES RESPONSE TO A FORMULATED PROBLEM OF ANY DEGREE OF SPECIFICITY.

10 PASAR CHARGE = $15 PROCESSING FEE + $2.25 PER MIN. COMPUTER TIME (AVE. = $40-$60, MAXM. = $95).

11 PADAT (*PA* DIRECT ACCESS TERMINAL), A SERVICE THAT ALLOWS REQUESTOR TO INTERACT WITH THE DATA BASE THROUGH A REMOTE COMPUTER TERMINAL IN HIS OWN FACILITY.

12 PADAT CHARGE = $25 SUBSCRIPTION FEE PER MO. + $1 PER MIN USE TIME (MINM. = 1 HR. USE TIME PER MO.).

13 PATELL (*PA* TAPE EDITION LEASE OR LICENSING), A SERVICE THAT MAKES *PA* MACHINE-READABLE RECORDS AVAILABLE TO INSTITUTIONS TO OPERATE THEIR OWN SEARCH AND RETRIEVAL SERVICES.

14 PATELL CHARGE = $3,000 FOR CURRENT YR., $2,000 FOR PREVIOUS YR., $800 FOR OTHER PREVIOUS YR.

For full information retrieval, write to:

PAIS
AMERICAN PSYCHOLOGICAL ASSOCIATION
1200 Seventeenth Street, N.W.
Washington, D.C. 20036

Those who would like to see PADAT in operation may request that a demonstration be arranged for their group at APA Central Office headquarters, or other negotiable locations.

Acc
Pocket &
Slip gone to the
Librarian
R Do
[illegible]

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 79, No. 6 November 1972

Articles

Hunger, Obesity, and the Ventromedial Hypothalamus **Richard E. Nisbett** 433

Foundations of an Attributional Theory of Performance **Andy Kukla** 454

The Theory of Cognitive Residues: A New View of Fantasy **David A. Rosenbaum** 471

Rhythmic (Hierarchical) versus Serial Structure in Speech and Other Behavior . . . . **James G. Martin** 487

Perception of Rotation in Depth: A Process Model **Myron L. Braunstein** 510

Relation between the Immediate Memory Span and the Memory Search Rate **J. Patrick Cavanaugh** 525

Theoretical Notes

Confidence Ratings in Recall: A Reanalysis **Douglas L. Hintzman** 531

Confidence Ratings for Individual Items in Recall **Harley A. Bernbach** 536

Gregory's Theory of Illusions: Some Disconfirming Evidence in the Case of the Muller-Lyer Illusion **Lester E. Krueger** 538

Comments on L. E. Krueger's "Disconfirming Evidence" of R. L. Gregory's Theory of Illusions . . . . . . **Richard L. Gregory** 540

Role of the Hypothalamus in Motivation: An Examination of Valenstein's Reexamination . . . . **Ernest H. Bergquist** 542

This is the last issue of Volume 79.
Title page and index for the volume appear herein.

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

J. W. Atkinson
Alan Baddeley
Harley A. Bernbach
Roy D'Andrade
R. H. Day
William K. Estes
C. R. Gallistel
R. Gelman
David M. Green
Harold W. Hake
Julian Hochberg
Bartley G. Hoebel
J. L. Jinks
Edward E. Jones
H. W. Leibowitz
A. M. Liberman
Eleanor Maccoby
James McGaugh
Bennett Murdock, Jr.
Ulric R. Neisser
Frank Restle
D. A. Riley
Sandra Scarr-Salapatek
R. L. Solomon
Saul Sternberg
A. J. Stunkard
Allan R. Wagner
Elizabeth Warrington

(A large group of psychologists has kindly agreed to assist in the editing of manuscripts submitted to the *Psychological Review*. In each issue the Editor will list in the *Review* those editorial consultants who have helped him process manuscripts for that issue, so that their efforts may be gratefully acknowledged.)

The *Psychological Review* is devoted to articles of theoretical significance to any area of scientific endeavor in psychology. Except for occasional articles, manuscripts exceeding 7,500 words will be discouraged. Ordinarily, manuscripts which consist primarily of original reports of research or of reviews of the literature should be submitted to other journals. Methodological and statistical articles are not appropriate for the *Psychological Review*. See the editorials in the *Review* (1965, **72**, 1–2, 1971, **78**, 1–2) for further details.

All manuscripts must be submitted in triplicate. In reference lists, give journal titles in full; do not abbreviate. All copy must be double spaced, including the references. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints; duplicate figures may be photocopies or pencil-drawn copies.

Beginning with the 1971 issue, Letters to the Editor will be published on an experimental basis. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obtained from the APA Central Office.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Manuscripts should be addressed to the Editor, George Mandler, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of materials copyrighted by the American Psychological Association: Written permission must be obtained for copying or reprinting any tables, figures, or text of more than 500 words in length. Persons or organizations who obtain permission to reproduce any such copyrighted materials must include on the first page of the reproduced materials the exact copyright notice which appeared on the Association's original publication. There will be a charge of $10.00 per page for such permission except when the requester is the author of the article to be reprinted, or when the article is to be reproduced in a limited number of copies solely for instructional purposes.

Harold P. Van Cott
Managing Editor

Lorraine Biros
Technical
Editor

Henry S. Salmon
Advertising and Business
Manager

Anne Redman
Subscription Manager

Published bimonthly by the
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Streets, Lancaster, Pa.
1200 Seventeenth St., N. W., Washington, D. C. 20036

$10.00 volume

$2.00 issue

Second-class postage paid at Lancaster, Pa.

Send all communications, including address changes, to 1200 Seventeenth St., N. W., Washington, D. C. 20036. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Diary No 727(6)
Date L.2
il- N L6
Bureau E-.l P.y: Research.

VOLUME 79, 1972

HAROLD P. VAN COTT
Managing Editor

LORRAINE BIROS
Technical Editor

HENRY S. SALMON
Advertising and Business Manager

ANNE REDMAN
Subscription Manager

Published Bimonthly
by The American Psychological Association, Inc.

Editorial Consultants

Henry Alker
Norman H. Anderson
Chris Argyris
J. W. Atkinson
Alan Baddeley
Albert Bandura
William F. Battig
Daryl Bem
Leonard Berkowitz
D. E. Berlyne
Harley A. Bernbach
Dalbir Bindra
Robert A. Bjork
A. H. Black
Donald S. Blough
Robert C. Bolles
Lyle E. Bourne
Gordon H. Bower
Jack W. Brehm
Peter E. Bryant
Donald T. Campbell
E. J. Capaldi
N. John Castellan
Michael Cole
R. G. Crowder
Roy D'Andrade
Robyn Dawes
R. H. Day
J. A. Deutsch
Morton Deutsch
Ebbe Ebbesen
Howard E. Egeth
Paul Ekman
William K. Estes
Seymour Feshbach
Leon Festinger
John M. Foley
Stephen S. Fox
J. L. Freedman
C. R. Gallistel
W. R. Garner
R. Gelman
Eleanor J. Gibson
R. Glaser
John Gibbon
James J. Gibson
Lewis R. Goldberg
J. A. Gray
B. F. Green
David M. Green
James Greeno
Richard L. Gregory
Robert G. Grice
Jeanne Gullahorn
Ralph N. Haber
Harold W. Hake
Kenneth R. Hammond
Charles S. Harris
James V. Hinrichs
Julian Hochberg
Bartley G. Hoebel
D. Howes
Earl B. Hunt
Leo M. Hurvich
Murray E. Jarvik
J. L. Jinks
N. Johnson
James Johnston
Edward E. Jones
Leon J. Kamin
Frederick H. Kanfer
Lawrence Karlin
H. Kelley
Howard H. Kendler
Tracy Kendler
G. Keppel
Walter Kintsch
Paul A. Kolers
Ronald L. Koteskey
David H. Krantz
T. K. Landauer
Richard S. Lazarus
H. W. Leibowitz
Samuel Leinhardt
Gerald S. Leventhal
Marvin Levine
Seymour Levine
A. M. Liberman
Elijah Lovejoy
R. Duncan Luce
Eleanor Maccoby
N. J. Mackintosh
John T. Macnamara
George F. Mahl
Edwin Martin
James McGaugh
Douglas Medin
Ronald Melzack
George A. Miller
Conrad G. Mueller
Bennett Murdock, Jr.
E. J. Murray
Jacob Nachmias
Ulric R. Neisser
Theodore M. Newcomb
Raymond S. Nickerson
Donald Norman
M. Frank Norman
Daniel Osherson
Seymour Papert
Karalyn Patterson
Albert Pepitone
M. Peterson
Irwin Pollack
Michael Posner
David Raab
P. M. A. Rabbitt
Robert A. Rescorla
Frank Restle
D. A. Riley
Milton Rokeach
Paul Rozin
Harris B. Savin
Sandra Scarr-Salapatek
Martin E. P. Seligman
Roger Shepard
Richard M. Shiffrin
R. L. Solomon
R. Sommer
George Sperling
Larry Squire
John E. R. Staddon
Larry Stein
Saul Sternberg
A. J. Stunkard
P. Suppes
N. S. Sutherland
James T. Tedeschi
John Theios
John Thibaut
Donald L. Thistlethwaite
Ewart Thomas
Richard F. Thompson
Thomas Trabasso
Milton A. Trapold
A. Treisman
Harry C. Triandis
Susan Ervin Tripp
Ledyard R. Tucker
R. E. Ulrich
Benton J. Underwood
Elliot S. Valenstein
P. Vitz
Victor H. Vroom
Allan R. Wagner
Hans Wallach
R. M. Warren
Elizabeth Warrington
Joachim F. Wohlwill
George Wolford
Ruth Wylie
J. I. Yellot, Jr.
Joseph L. Young

CONTENTS OF VOLUME 79

Articles

Anderson, Daniel R. See Shepp, Bryan E.

Anderson, John R., and Bower, Gordon H. **Recognition and Retrieval Processes in Free Recall** 97

Bergquist, Ernest H. **Role of the Hypothalamus in Motivation: An Examination of Valenstein's Reexamination** 542

Bennett, Edward L. See Diamond, Marian C.

Bernbach, Harley A. **Confidence Ratings for Individual Items in Recall** 536

Biederman, Gerald B. **Continuity Theory Revisited: Reply to D. B. Berch** 178

Bolles, Robert C. **Reinforcement, Expectancy, and Learning** 394

Bower, Gordon H. See Anderson, John R.

Braunstein, Myron L. **Perception of Rotation in Depth: A Process Model** 510

Campbell, Byron A., and Spear, Norman E. **Ontogeny of Memory** 215

Cavanaugh, J. Patrick. **Relation between the Immediate Memory Span and the Memory Search Rate** 525

Clark, Herbert H. **On the Evidence Concerning J. Huttenlocher and E. T. Higgins' Theory of Reasoning: A Second Reply** 428

Coren, Stanley. **Subjective Contours and Apparent Depth** 359

Devoe, Shannon. See Wiener, Morton.

Diamond, Marian C. See Rosenzweig, Mark R.

Dienstbier, Richard A. **A Modified Belief Theory of Prejudice Emphasizing the Mutual Causality of Racial Prejudice and Anticipated Belief Differences** 146

Geller, Jesse. See Wiener, Morton.

Gibbon, John. **Timing and Discrimination of Shock Density in Avoidance** 68

Green, David M. See Luce, R. Duncan.

Greeno, James G. See Martin, Edwin.

Gregory, Richard L. **Comments on L. E. Krueger's "Disconfirming Evidence" of R. L. Gregory's Theory of Illusions** 540

Groen, Guy J., and Parkman, John M. **A Chronometric Analysis of Simple Addition** 329

Groves, Philip M., and Lynch, Gary S. **Mechanisms of Habituation in the Brain Stem** 237

Hammond, Kenneth R., and Summers, David A. **Cognitive Control** 58

Higgins, E. Tory. See Huttenlocher, Janellen.

Hintzman, Douglas L. **On Testing the Independence of Associations** 261

Hintzman, Douglas L. **Confidence Ratings in Recall: A Reanalysis** 531

Huttenlocher, Janellen, and Higgins, E. Tory. **On Reasoning, Congruence, and Other Matters** 420

Kemler, Deborah G. See Shepp, Bryan E.

Koteskey, Ronald L. **A Stimulus-Sampling Model of the Partial Reinforcement Effect** 161

Krebs, Marjorie J. See Teichner, Warren H.

Krueger, Lester E. **Gregory's Theory of Illusions: Some Disconfirming Evidence in the Case of the Muller-Lyer Illusion** 538

Kukla, Andy. **Foundations of an Attributional Theory of Performance** 454

Lockhead, G. R. **Processing Dimensional Stimuli: A Note** 410

Luce, R. Duncan, and Green, David M. **A Neural Timing Theory for Response Times and the Psychophysics of Intensity** 14

Lynch, Gary S. See Groves, Philip M.

Macnamara, John. **Cognitive Basis of Language Learning in Infants** 1

Martin, Edwin, and Greeno, James G. **Independence of Associations Tested: A Reply to D. L. Hintzman** 265

Martin, James G. **Rhythmic (Hierarchical) versus Serial Structure in Speech and Other Behavior** 487

Massaro, Dominic W. **Preperceptual Images, Processing Time, and Perceptual Units in Auditory Perception** 124

Møllgaard, Kjled. See Rosenzweig, Mark R.

Moore, Dewey J. **A Response to R. A. Webb's Comments** 280

NISBETT, RICHARD E. **Hunger, Obesity, and the Ventromedial Hypothalamus** 433
PARKMAN, JOHN M. See GROEN, GUY J.
PARLEE, MARY B. **Comments on "Roles of Activation and Inhibition in Sex Differences in Cognitive Abilities" by D. M. Broverman, E. L. Klaiber, Y. Kobayshi, and W. Vogel** 180
ROSENBAUM, DAVID A. **The Theory of Cognitive Residues: A New View of Fantasy** 471
ROSENZWEIG, MARK R., MØLLGAARD, KJLED, DIAMOND, MARIAN C., AND BENNETT, EDWARD L. **Negative as well as Positive Synaptic Changes May Store Memory** 93
RUNBINOW, STUART. See WIENER, MORTON.
SHALLICE, TIM. **Dual Functions of Consciousness** 383
SHEPP, BRYAN E., KEMLER, DEBORAH G., AND ANDERSON, DANIEL R. **Selective Attention and the Breadth of Learning: An Extension of the One-Look Model** 317
SIEGEL, JANE A., AND SIEGEL, WILLIAM. **Absolute Judgment and Paired-Associate Learning: Kissing Cousins or Identical Twins?** 300
SIEGEL, WILLIAM. See SIEGEL, JANE A.
SIMON, HERBERT A. **Complexity and the Representation of Patterned Sequences of Symbols** 369
SMITH, PETER G. See THEIOS, JOHN.
SPEAR, NORMAN E. See CAMPBELL, BYRON A.
STOKOLS, DANIEL. **On the Distinction between Density and Crowding: Some Implications for Future Research** 275
SUMMERS, DAVID A. See HAMMOND, KENNETH R.
TEICHNER, WARREN H., AND KREBS, MARJORIE J. **Laws of the Simple Visual Reaction Time** 344
THEIOS, JOHN, AND SMITH, PETER G. **Can a Two-State Model Account for Two-Choice Reaction-Time Data?** 172
TVERSKY, AMOS. **Elimination by Aspects: A Theory of Choice** 281
VIKIS-FREIBERGS, VAIRA. **A Stochastic Model for Free Association Response Hierarchies?** 268
WALLSTEN, THOMAS S. **Conjoint-Measurement Framework for the Study of Probabilistic Information Processing** 245
WEBB, ROGER A. **A Comment on D. J. Moore and D. A. Shiek's "Toward a Theory of Early Infantile Autism"** 278
WIENER, MORTON, DEVOE, SHANNON, RUNBINOW, STUART, AND GELLER, JESSE. **Nonverbal Behavior and Nonverbal Communication** 185

LETTER

POWER, RODERICK P. **A Possible Test of Hebb's Hypothesis Concerning Imagery** 368
HEBB, DONALD O. **Reply to "A Possible Test of Hebb's Hypothesis Concerning Imagery"** 368

PSYCHOLOGICAL
REVIEW

VOL. 79, NO. 6 NOVEMBER 1972

HUNGER, OBESITY, AND THE VENTROMEDIAL HYPOTHALAMUS[1]

RICHARD E. NISBETT[2]

University of Michigan

There are striking behavioral parallels between obese individuals and hungry individuals, suggesting that many obese individuals are actually in a chronic state of energy deficit and are genuinely hungry, perhaps because they attempt to hold their weight below its biologically dictated "set point." This conclusion is consistent with the remarkable similarities between the behavior of obese human beings and the behavior of animals that become obese after lesions of the ventromedial hypothalamus. Consideration of these parallels suggests a reexamination of traditional views of hunger and the ventromedial hypothalamus syndrome.

Obesity has for many years been attributed to gluttony. The obese individual is pronounced by society, and often by his physician, to be lacking in will power. If one excludes the constellation of obesities that involve obvious glandular and hormonal causes (Gordon, 1970), obesity is without any apparent primary etiology. Almost all the major physiological characteristics of "simple" obesity can be explained as consequences of the increased adipose tissue stores. The cause for the increase in fat stores has remained elusive, and this lack of any obvious cause seems to have reinforced the conclusion that the overweight individual is a glutton.

A different explanation of obesity has recently gained plausibility. Simply put, this view suggests that some individuals have no choice but to be fat. Like all biological parameters, the amount of body fat stores varies from individual to individual, and some individuals have a higher base-line level of adipose tissue than others. They are biologically programmed to be fat.

Physiological Basis of Obesity

This idea of an elevated base line of adipose tissue stores has recently been given a quite explicit physiological meaning. Body fat in the form of esterified fatty acids is stored in specialized fat cells called adipocytes or lipocytes. The quanitity of fat an individual has is thus a function of the number and size of the adipocytes he possesses. In an effort to determine whether obese in-

---

[1] The author is deeply grateful to Terry Powley for extensive criticism and assistance. Thanks are also due to Nicholas Mrosovsky, Judith Rodin. L. Reid Hanson, and Katerina Borer for criticism of an early draft, and to Jules Hirsch for advice and assistance on some of the research. This research was supported by National Science Foundation Grant GB 33918.

[2] Requests for reprints should be sent to Richard E. Nisbett, Research Center for Group Dynamics, Institute for Social Research, University of Michigan. Ann Arbor, Michigan 48106.

© 1972 by the American Psychological Association, Inc.

dividuals have more fat cells or larger fat cells, Hirsch and Knittle and their colleagues (Hirsch & Han, 1969; Hirsch & Knittle, 1970; Knittle & Hirsch, 1968) and Bjorntorp and his co-workers (1972) have counted and measured the size of adipocytes in normal and obese subjects. Both laboratories report that adiposity of obese humans is determined primarily by the number, as opposed to size, of fat cells. While there is often a modest increase in the size of adipocytes (hypertrophy), simple obesity seems to be basically an increase in the number of cells (hyperplasia). In one sample, obese subjects were found to have three times as many fat cells as the normal weight group (Knittle & Hirsch, 1968).

The importance of these data for an understanding of obesity becomes clear when it is realized that the number of fat cells an adult possesses is essentially fixed and stable. Mature adipocytes are considered to be x nonmitotic cells (Goss, 1966), and thus the number of fat cells would be stable in the adult organism. A period of severe starvation beginning as early as the sixth week of life in rats is incapable of reducing the number or permanently altering the size of fat cells (Hirsch & Han, 1969). Starvation and dieting in adult humans decrease the size but not the number of adipocytes to any appreciable extent (Hirsch & Knittle, 1970). After a period of weight loss the ex-obese individual retains the same excessive number of fat cells to be filled back up the moment will power fades. Conversely, overeating in the adult does not stimulate the growth of more adipocytes. Hirsch and Han (1969) have found that in adult animals lesions of the ventromedial nucleus of the hypothalamus that result in hyperphagia and weight gain cause a drastic increase in the size of fat cells but not an increase in number. Prison volunteers paid to get fat similarly increase the size but not the number of fat cells (Sims, Kelleher, Horton, Gluck, Goldman, & Rowe, 1968). Thus it appears that a person with a given number of fat cells is not able to decrease that number through dieting or to increase that number through bad habits or overeating. The individual with a large number of fat cells will have a higher base line of body fat. That is, he will be constitutionally fat.

Further, the two variables that seem to affect the final base line of adipose tissue stores cannot be readily manipulated. The first factor that seems of primary importance in establishing the organism's level of adipose tissue is heredity. That adiposity has a strong genetic component is indicated by research on animal fat deposition. Several investigators have shown that strains of rats vary considerably in their genetic potential to accumulate fat stores on various diets (Schemmel, Mickelsen, & Gill, 1970; Schemmel, Mickelsen, & Tolgay, 1969; Marshall, Smith, Munson, & Lehmann, 1969). While human data are always open to interpretation on environmental grounds, the magnitude of the association between overweight in parents and their children makes it seem likely that there is a strong hereditary component in human obesity as well. For example, Davenport (1923) reported that in a sample of 51 children all born of parents who were slender, none of the children were of more than average weight and the great majority were themselves slender. In a sample of 37 children with very obese parents, none were slender, all were of at least average weight, and one-third were themselves very obese.

The second factor that seems to be critical in establishing the base line of adipose tissue stores is early nutritional experience. Kennedy (1957) and Widdowson and McCance (1960) have shown that the percentage of fat in the bodies of rats is permanently affected by caloric intake during the suckling period. Knittle and Hirsch (1968), looking at the cellularity of the fat pads in rats exposed to this same kind of infantile nutritional experience, have shown that early nutritional experiences can effect permanent changes in cell number and size of the fat depots. Knittle and Hirsch also point out that extreme human obesity seems to involve a pattern of cellularity and metabolism of adipocytes similar to that seen in rats subjected to early "overnutrition." Thus the number of adipocytes an adult possesses is probably influenced both by the hereditary and early nutritional history of the individual.

The proposition that there are individual differences in the base line for fat stores becomes important in view of the fact that recent evidence suggests that the adipose tissue mass is defended by the central nervous system. The classic hypothalamic feeding centers appear to adjust food intake so as to maintain fat stores at the base line or "set point" level. Evidence for this view comes from studies of the effects of destruction of these centers. Hoebel and Teitelbaum (1966) have convincingly demonstrated that animals with damage to the ventromedial nucleus of the hypothalamus regulate their weight at a new, higher level. A number of investigators have indicated that weight increases resulting from such ventromedial hypothalamus lesions reflect exclusively an elevated set point for adipose tissue; the lesions do not affect the amount of nonfat body tissue. Reciprocal results have been described for the lateral hypothalamic area. Powley and Keesey (1970) showed that animals with lateral hypothalamic lesions lower the level at which they regulate their body weight. Powley (1971) has indicated that this lower weight level again reflects a change in the set point for adipose tissue stores exclusively. The evidence thus suggests that in its intact state, the ventromedial hypothalamus acts as a brake to keep an animal from adding excess fat, a proposal first made by Kennedy (1957). A similar description now seems appropriate for the lateral hypothalamus: in its intact state it serves to keep the animal from losing fat tissue. As Hervey (1969) has pointed out, this regulation by the hypothalamic centers is exquisitely precise. An average daily error of so much as a handful of peanuts would be sufficient to add 10 pounds of adipose tissue per year. Obviously, most people regulate their weight within a much finer tolerance than this.

If it is correct that the hypothalamic centers defend adipose tissue set point, then there seems to be no reason to assume that they defend the same set point in all individuals of the same height and bone structure. Instead, it is possible that the *hypothalamic centers defend different baselines in different individuals,* maintaining whatever set point has been established by heredity and by nutritional conditions during the malleable period of childhood. This proposition suggests a rethinking of much of the behavioral data on obesity. If correct, it would indicate that obesity for some individuals would represent a "normal" or "ideal" body composition. The use of such terms as "pathology" or "disease" to account for obesity would be inappropriate in such cases. Further, and quite paradoxically, a consideration of the sociology of obesity suggests that many individuals in an "*over*weight" population will actually be "*under*weight." The person with an elevated adipose tissue set point will be under considerable social (and often medical) pressure to reduce his weight. Given, once again, that obesity or an elevated level of fat stores is his natural state, such weight reducing would hold him below his physiologically dictated normal body weight. The central nervous system of such a person should respond to weight reduction just as would the central nervous system of a leaner individual, that is, with symptoms of energy deficit. Such a person would literally be starving all the time.

There are in fact some striking parallels between the behavior of obese humans and that of hungry organisms. In elucidating these parallels, it will be helpful to review first the behavior changes that are induced by hunger. Some of these changes are familiar, but many of them are not at all widely known.

THE BEHAVIOR OF HUNGRY ORGANISMS

Eating Behavior

Many of the effects of hunger on eating behavior are obvious to anyone who has ever been hugry. The food-deprived organism, given an opportunity to eat, eats more and eats more rapidly than the less deprived organism; the food-deprived organism is more likely to eat in a novel environment containing food than the less deprived organism, etc. Less obvious are the effects of hunger on response to the taste of food. Until recently, the effect of hunger on taste responsiveness has not been examined, probably because it has been taken for granted that the hungry organism is undiscriminating, eating large quantities of any available food without regard to its taste properties. This assumption now seems to be quite wrong. The evidence suggests that the deprived organism consumes proportionally more good-tasting food than does the less deprived organism.

The first indication that the hungry animal is highly taste responsive came from studies examining the effects of food deprivation on preference for sweet solutions over water. Bacon, Snyder, and Hulse (1962) found that the food-deprived rat, given access to both water and water flavored with saccharin, shows a much more marked preference for the sweet solution than the nondeprived animal. Valenstein (1967) has shown that when nondeprived animals are offered a choice between highly sweet saccharin solution and less sweet glucose solution, they prefer the less sweet glucose solution. When animals are deprived of food, however, they come to prefer the more sweet though nonnutritive, saccharin solution. Williams and Campbell (1961) deprived rats of food to varying degrees and then allowed them access for 24 hours to either milk or quinine-adulterated milk. With increasing hunger, larger amounts both of milk and of quinine-adulterated milk were consumed. But, up to very extreme degrees of deprivation, the difference between the amount of milk consumed compared to the amount of adulterated milk consumed increased steadily with increasing deprivation. Jacobs and Sharma (1969) offered dogs and rats either standard laboratory chow or chow with improved flavor due to the addition of fats or saccharin. Animals were either allowed to eat whenever they wished or were given only a brief meal (15 minutes to 2 hours) in every 24 hours. Ad-lib and deprived animals consumed equivalent amounts of standard chow, but deprived animals consumed much more good tasting chow than ad-lib animals. The evidence thus suggests that the deprived animal becomes increasingly more responsive to taste, consuming proportionally more and more good-tasting food, compared to standard fare, as deprivation increases.

Jacobs and Sharma (1969) have gone further than this, suggesting that the deprived animal is not only relatively more accepting of good food but also *rejects bad food* relatively more. In one of their studies, deprived and ad-lib dogs were offered access only to cellulose-adulterated chow. The intake of deprived animals, offered only one meal a day, remained at one-third ad-lib levels for a period of *10 days*. Deprived animals were physically capable of consuming much more than that during the brie meal, since when food was diluted mucl more, but with water (which makes a more pleasant-tasting diluent), deprived animals ate as much during their brief meal as ad-lib animals did over a 24-hour period. Additional supportive data comes from work by Gross (1968), who subjected young rats, for a period of weeks, to deprivation schedules which were so severe that they could not maintain normal weight. When returned to ad-lib schedules, these still-hungry animals rejected cellulose-diluted chow relatively more than did animals always maintained on ad-lib schedules. (And, as would be expected from the literature concerning response to good-tasting food, the formerly deprived animals consumed relatively more chow to which fats had been added than did the always ad-lib animals.)

It should be noted that Jacobs and Sharma's (1969) proposition appears not to hold in its strongest form. When animals are given identical amounts of time to eat and are totally starved for varying amounts of time before the opportunity to eat, it appears to be the case that intake of even very bad-tasting food is positively related to number of hours of starvation (Hoff & Gentile, 1969; Tenen & Miller, 1964; Williams & Campbell, 1961). It is only when opportunities to eat are of varying lengths (Jacobs & Sharma, 1969) or deprivation is achieved by some technique other than total starvation (Gross, 1968; Nisbett, Hanson, Harris, & Stair, 1972) that the finickiness of the [illegible] prived animal is observed. T[illegible] the rebel on commonsen[illegible] adult possesses is [illegible] limited finickiness, it should be pointed out that there might be a survival advantage for animals who, when hungry, are chary of bad-tasting (and possibly poisonous or injurious) food.

It thus appears that the food-deprived organism is highly discriminating. Deprivation enhances the acceptance of good-tasting food and, under at least some circumstances, leads to the rejection of bad-tasting food. Hunger also, as will be seen, has powerful effects on behavior not directly related to eating.

Emotionality, Activity, and Sex

Most of what we know about the effects of extreme hunger on general behavior comes from the classic work of Keys, Brožek, Henschel, Mickelsen, and Taylor (1950) on the effects of semistarvation on human beings. The work of Keys et al. served to verify observations that had long been made of humans in famine and concentration camp situations (e.g., Leyton, 1946). Their conclusions have also been essentially duplicated by later workers (Brown & Pulsifer, 1965; Meerloo & Klauber, 1952). The subjects employed by Keys et al. were conscientious objectors who volunteered to lose 24% of their body weight over a 24-week period. This was accomplished by restricting daily caloric intake to the vicinity of 1,600 calories while maintaining normal work routines.

Keys et al. (1950) studied an extremely wide range of behaviors. Particularly marked changes were observed in three distinct areas. The men became, during the course of their fast, progressively more prone to emotional upset, more apathetic and inactive, and less interested in sex.

Emotionality. The emotional effects of hunger have long been known. An observer of the 1918–1922 famine in Russia coined the term "neurosis of inanition" to describe the effects of hunger, which in children, "manifested itself in a combination of unusual vindictiveness, extraordinary irritability, and chronic crying . . . (cited in Keys et al., p. 909)." In the subjects studied seenKeys et al., irritability was quite marked defend the sa[illegible] temper so frequent that the of the same height an[illegible] during the control period had to be curtailed. Interestingly, periods of elation also occurred (inevitably followed by periods of depression), which the men attributed to a "starvation high." At a more formal level of analysis, Minnesota Multiphasic Personality Inventory scales were administered at periodic intervals. Over the course of semistarvation, most of the scales increased in elevation, with the effects being particularly marked on the so-called "neuratic triad" of depression, hypochondriasis, and hysteria.

Activity. The men became progressively more apathetic and inactive during the course of semistarvation. They did only what was necessary, and that only with an effort of will. This lassitude, together with a striking lack of sexual interest, is well conveyed by a quotation from a diary kept by one of the men.

> I am one of about three or four (out of 32) who still go out with girls. I fell in love with a girl during the control period but I see her only occasionally now. It's almost too much trouble to see her even if she visits me at the lab. It requires effort to hold her hand. Entertainment must be tame. If we see a show, the most interesting part of it is contained in scenes where people are eating . . . love scenes are completely dull [Keys et al., 1950, p. 853].

It should be noted parenthetically that the activity level of rats actually increases with increasing food deprivation. This diametrically opposite effect probably has to do with the heat-dissipation problems of small mammals. When the energy source provided by adipose tissue begins to dwindle, rodents appear to have difficulty in maintaining body temperature (Stevenson & Rixon, 1957). Activity makes up in part for energy store losses. Hamilton (1969) has shown that the deprived rat will work to turn on a heat source, and at marked degrees of deprivation prefers this source of heat to increased body activity. Stevenson and Rixon have shown that when ambient temperatures are relatively high, the increase in activity due to food deprivation is relatively small.

Sex. Semistarvation all but destroyed sexual motivation. Several engagements were broken, although as Keys et al. (1950) noted, the emotional disorders and lack of activity drive may have played a role in this. Masturbation and nocturnal emissions ceased or were greatly reduced. It seems likely that the hyposexuality of hunger is mediated at least in part by a drop in gonadal hormones. Keys et al. found a one-third decrease in urinary excretion of 17 ketosteriods, which covary with testosterone level. Miller, Mickelsen, and Keys (1948) found a two-thirds drop in 17 ketosteroids in a similar population undergoing total abstinence from food for four days.

THE BEHAVIOR OF OBESE HUMANS

It is fair to say that the major areas in which obese humans have been shown to differ behaviorally from normal weight humans parallel almost exactly the areas affected by hunger. This is true despite the fact that almost none of the research on obese humans seems to have been guided by the hypothesis that the obese are hungry. On the contrary, the opposite assumption has often been the case.

Eating Behavior

Most of the pertinent empirical work on the eating behavior of obese humans stems from a program initiated by S. Schachter (1967, 1968, 1971). In the course of that program, it has been discovered that like the hungry organism, the obese human eats more food at a given sitting and eats more rapidly (Schachter, 1971), and has a greater readiness to eat as indicated by the fact that in a novel environment containing food the obese individual is more likely to eat than the normal individual (Ross, 1969). Like the hungry organism, the obese individual is highly taste responsive. Nisbett (1968a, in press) has shown that the difference between the amount of liked and disliked food eaten by obese subjects is greater than for normal weight subjects. Hashim and Van Itallie (1965) have shown that obese subjects greatly restrict caloric intake when allowed access only to bland, unappetizing liquid formula. Normal individuals subjected to the same dietary regime maintain normal caloric intake. Decke (cited in Schachter, 1971) gave milk shakes or quinine-adulterated milk shakes to obese and normal subjects. Obese subjects consumed more unadulterated milk shake and less adulterated milk shake than normal subjects.

A peculiarity of the eating behavior of the obese is that it seems to reflect a constant, moderately strong degree of hunger, which is remarkably little affected by conditions that increase or lower the normal individual's interest in food. Schachter, Goldman, and Gordon (1968) asked obese and normal subjects to "taste" a variety of crackers. Half of their subjects had just eaten two roast beef sandwiches and half were in a state of mild food deprivation. Normal subjects, as would be expected, ate fewer crackers in the "tasting session" if they had just eaten two sandwiches than they did if deprived. Overweight subjects, however, ate fully as many crackers after eating the sandwiches as they did if deprived. Similarly, Nisbett (1968a) found that self-report of hunger varied with deprivation state for normal subjects but did not vary for obese subjects. And Nisbett and Kanouse (1969) have shown that while normal weight individuals do more and more impulse buying in supermarkets as their state of food deprivation increases, overweight shoppers do a moderate amount of impulse buying regardless of their state of deprivation.

Cabanac and Duclaux (1970) have combined into one experimental paradigm a demonstration both of exaggerated responsiveness to taste cues and of unresponsiveness to short-term changes in nutritional state. Normal individuals find otherwise pleasant-tasting sweet solutions to be aversive when they sample them a few minutes after a gastric load of glucose. Obese individuals find the solutions somewhat more pleasant tasting than do normal individuals in the absence of a glucose load, and the glucose load decreases their attractiveness for the obese individual hardly at all.

The insensitivity of the obese to short-term nutritional change is reminescent of an observation reported by Keys et al. (1950). During the rehabilitation period following semistarvation, their subjects were allowed to eat all they wanted. It was not uncommon for the men to eat to the point where stomach distention made it impossible to continue and yet report that they still felt hungry and desirous of eating more.

The similarities between the eating behavior of obese humans and that of hungry organisms are therefore quite marked. The similarities in other behavioral areas are almost equally pronounced.

Emotionality, Activity, and Sex

Emotionality. By a variety of indicators, overweight humans appear to be more emotional than normal weight individuals. Schachter et al. (1968) found that overweight subjects reported themselves to be more frightened at the prospect of receiving electric shocks than did normal subjects, and Rodin (1972) found that overweight subjects worked harder to avoid shock than normal subjects. Rodin, Elman, and Schachter (1972) found that the proofreading and monitoring performance of obese subjects deteriorated when they listened to emotionally charged tapes. Performance of normal subjects did not deteriorate. The greater emotionality of the obese is reflected in measures of emotional adjustment in daily life. Moore, Stunkard, and Srole (1962) have analyzed the mental health data collected from a random sample of the residents of midtown Manhattan (Srole, Langner, Michael, Opler, & Rennie, 1962). The obese respondents were found to obtain poorer scores on virtually all the tests of emotional adjustment. An implication of the present view is that obese individuals should not be more emotionally disturbed unless their weight is below its high set point. Indirect evidence that this is the case comes from a consideration of the social class data presented by Moore et al. The lower socioeconomic (SES) sample was much more obese than the middle and upper SES samples. Perhaps there is less concern with overweight on the part of lower SES individuals, or perhaps medical or social pressure to lose weight is less for that group. If so, then a higher proportion of lower SES individuals should be at set point, and the relationship between overweight and emotional disturbance should be weaker for that group. At any rate, that was the case. A reanalysis by the present writer of the mental health data of Moore et al. shows that it is *only* among the middle- and upper-classes that obesity is associated with increased emotional disorder.

A second implication of the present view is that the emotional state of obese individuals should deteriorate still further when they attempt to lose large amounts of weight. Although the point is in dispute, a number of investigators reported that this is the case (notably, Bruch, 1957; Gluckman & Hirsch, 1968; Gluckman, Hirsch, McGully, Barron, & Knittle, 1968; Stunkard, 1957). The most commonly reported symptoms include depression and irritability but reports of more serious effects, including psychosis and suicide attempts, are not uncommon.

Activity. Overweight individuals are markedly less active than normal weight individuals. Chirico and Stunkard (1960) gave pedometers to obese and normal subjects, matched for occupational role, and measured the amount they walked in the course of a day. In the great majority of cases, the normal subject outwalked the obese subject. Bullen, Reed, and Mayer (1964) took motion pictures of girls engaged in sports and found large differences in the amount of time spent in strenuous activity. Obese girls only very rarely engaged in such activity. Bullen et al. and a number of subsequent investigators questioned overweight and normal subjects about their daily routines. Overweight individuals report that a considerably higher fraction of their time is spent in the most sedentary type of activities.

Sex. The data are least complete in the area of sexual behavior. The evidence that does exist is consistent with the parallels to the hungry organism. Bruch (1957) reported that her obese patients as a group had remarkably little sexual interest. Results from a preliminary study by R. E. Nisbett and J. Platt (unpublished data) suggest that overweight college males have less sexual motivation than do normal weight males. This tendency was quite marked for the incidence of nocturnal orgasms, an aspect of sexual behavior that would seem to be little affected by social circumstances.

OBESITY, STARVATION, AND FREE FATTY ACIDS

Clearly, then, in a variety of behavior areas, the obese individual and the hungry individual have much in common. Both appear more taste responsive, less responsive to postingestional feeding cues, more emotional, less active, and hyposexual. These parallels raise the question of whether there are also physiological indications that the obese are hungry.

There are at the present time no universally accepted physiological indicators of energy deficit. Probably, however, the most generally agreed-upon index is the level of free, unesterified fatty acids in the blood. When the organism is food deprived, is cold, or exercises, free fatty acids are mobilized from adipose tissue to meet the increased energy demand (Gollnick, 1967; Gordon & Cherkes, 1956; Hanson, Johnson, & Zaharto, 1965; Mager & Iampietro, 1966; Van Itallie & Hashim, 1960). Starvation reliably elicits elevated free fatty acid levels (Walker & Remley, 1969). When the organism eats, there is a rapid fall of free fatty acid level.

A large number of recent studies indicate that free fatty acid level is higher in obese individuals than in individuals of normal weight (Björntorp, Bergman, & Varnouskas, 1969; Björntorp, Bergman, Varnouskas, & Lindholm, 1969; Dole, 1956; Gordon, 1960; Heald, Mueller, & Dougela, 1965; Issekutz, Bortz, Miller, & Wroldsen, 1967; Jahnke, Gries, Wallenfels, & Schulte, 1964; Leibermeister, Daweke, Gries, Schilling, Grüneklee, Probst, & Jahnke, 1968; Morse & Mahabir, 1965). Until quite recently, this elevation of free fatty acid levels in the obese was considered to be simply a consequence of having a large adipose tissue depot. It now seems quite clear that this is not the case, however, for several reasons: (*a*) Experimentally induced obesity, both in men (Sims, Kelleher, Horton, Gluck, Goldman, & Rowe, 1968) and rats (Steffens, 1969), results in slightly decreased free fatty acid levels. (*b*) In one of the few cases in the literature of a failure to find significantly elevated free fatty acid levels (Rabinowitz & Zierler, 1962), the authors explicitly noted that all their obese subjects had just gone through a period of several months' active weight gain. Moreover, Gordon (1960) has reported that in obesity "clearly related to overeating," the elevation of free fatty acid levels was less marked or altogether absent. (*c*) When obese individuals lose weight, free fatty acid levels increase still further (Genuth, 1966; Issekutz et al., 1967; Leibermeister et al., 1968).

Parenthetically, it may be noted that perhaps the chief reservation about accepting free fatty acid level as an index of hunger has been the fact that most overweight persons show such an elevation. If there is validity to the thesis that many obese individuals are below their set point for weight, this would clearly explain the discrepancy concerning the interpretation of free fatty acid levels.

An equally important fact for an understanding of obesity and hunger is that free fatty acid level in the obese is relatively inflexible. It does not respond as does free fatty acid level in normals to changes in deprivation state. Several investigators have reported that during a day's fast, free fatty acid levels of the obese increase only slightly or not at all (Corvilain, Loeb, Champenois, & Abramow, 1961; Gordon, 1960; Heald et al., 1965; Opie & Walfish, 1963; Morse & Mahabir, 1964). Normals, in contrast, start with lower free fatty acid levels than the obese, and in the vicinity of 20–24 hours deprivation, equal obese free fatty acid levels. The free fatty acid response of the obese to a gastric load of glucose is similarly deficient. During a fast, the free fatty acid levels of normal subjects plummet rapidly to nonfasting levels in response to oral glucose. The free fatty acid response of obese subjects to oral glucose is both slower and less complete than in normal subjects (Bagdade, Porte, & Bierman, 1969; Gordon, 1970; Leibermeister et al., 1968; Morse & Mahabir, 1964).

The importance of these data for an understanding of obesity seems very great. It is apparent that the behavioral and physiological data are in accord. Overweight individuals behave as if they were always—and inflexibly—hungry. They eat more per meal, they eat more rapidly, they are more responsive to taste; and neither their eating behavior, their self-report of hunger, nor even the attractiveness of supermarket food is much affected by short-term changes in deprivation state. The physiological evidence fully justifies this pattern of behavior: free fatty acid level is always—and inflexibly—high.

RECENT SUPPORTIVE EVIDENCE

To review the present line of reasoning, if the central nervous system operates so as to defend different set points in different individuals of the same height and bone structure, rather than by defending in all individuals the most esthetic and healthful "ideal" set point, then a quite paradoxical situation could result: an individual could be obese and yet be at or even below the set point which his system is trying to defend.

If there is some truth to this notion, then it is quite possible that a great many obese persons, perhaps even the majority in our society, exist at a weight level below their biologically dictated set points. Social pressure against overweight is great enough in modern Western societies to insure that many overweight individuals are probably trying most of the time to curtail their food intake and lower their weight. The results would be energy deficit, high free fatty acid level, and hunger.

If obese individuals have higher set points but try to stay below them, and if they have elevated free fatty acid levels and a "hungry" pattern of behavior because of their attempts, then several things should follow: (*a*) When anyone, obese or normal, is below his set point for weight, his eating behavior should resemble that of obese individuals. (*b*) When obese individuals are at their set points, their behavior should resemble that of normal individuals. (*c*) No association should be found between degree of overweight and hunger-like behaviors in *nonhumans*, since there is no social pressure operating on overweight animals to lose weight.

Some support can be marshalled for each of these expectations. Cabanac, Duclaux, and Spector (1971), whose line of reasoning is similar to the position presented here and has influenced it greatly, have examined the effects of weight loss on taste responsiveness. In their study of the affective response to sweet solutions before and after gastric glucose loads, Cabanac and Duclaux (1970) had observed that their two *most obese* subjects gave a "normal" response to glucose, that is, they disliked the sweet solutions after swallowing the glucose load (personal communication, 1970). On the other hand, the most *underweight* of their subjects gave an "obese" response, that is, did not dislike the sweet solutions after the glucose load. These observations suggested that the very most obese subjects had ignored social pressure and were at or perhaps above their set points. The most underweight of the normal subjects was, perhaps because of illness or emotional stress, below his set point. The remaining overweight subjects would have been below set point, and hence hungry and unresponsive to the glucose load. The remaining normal subjects would have been at set point, not hungry, and thus responsive to the glucose load. In order to test the hypothesis that weight reduction would result in a failure to respond to glucose loads, Cabanac et al. (1971) studied reactions to glucose loads before, during, and after a period of weight loss. The authors studied their own responses over the course of a 10% weight loss. Prior to loss, and after weight was regained, glucose loads made sweet solutions unpleasant to the taste. During the loss and during maintenance at the low weight, glucose loads had no effect on the pleasantness of the sweet solutions.

Cabanac et al. (1971) interpreted their results as evidence for the existence of a "ponderostat," a mechanism capable of detecting whether the organism is at its set point for weight. When the organism is below its set point, the ponderostat detects this fact, with the result that events which would signal satiety to the organism at set point are ignored. Thus, Cabanac and Duclaux' (1970) obese subjects, except for the very fattest, would have been below their set points because of responsiveness to social pressure and hence immune to postingestional cues indicating a transitory, but insufficient, improvement in energy balance. This

TABLE 1

GRAMS OF ICE CREAM EATEN AS A FUNCTION OF ITS TASTE, SUBJECTS' WEIGHT, AND SUBJECTS' ESTIMATED NEARNESS TO SET POINT

Probably near set point

Taste	Underweight (deviation > −10)	Normal (no history of overweight)	Overweight (deviation > +40)
Good	138.7	136.9	137.2
n	19	16	4
Bad	69.5	38.6	64.6
n	18	19	6
Good–bad	69.2	98.3	72.6

Probably below set point

	Underweight (deviation ≤ −10)	Normal (history of overweight)	Overweight (deviation ≤ +40)
Good	152.2	198.3	245.9
n	9	12	24
Bad	47.0	18.1	58.5
n	10	9	22
Good–bad	105.2	180.2	187.4

suggestion, which is entirely congenial to the present line of reasoning, implies that a reanalysis of previous eating behavior data would reveal that the most extremely obese subjects would show a "normal" pattern and the most underweight would show an "obese" pattern. An experiment by Nisbett (1968a) allows a fair test of this expectation. The responses to good- and bad-tasting ice cream of a very large number of underweight, normal, and overweight subjects, 56 in each category, was studied. In addition, the normal category included a large number of formerly overweight subjects who had been invited to the experiment on the basis of weight records indicating that they were overweight but who lost sufficient weight in the interim to qualify as being of normal weight. These latter subjects would be expected to be below their set points and should show the obese pattern.

Amounts of good and bad (quinine-adulterated) ice cream eaten, as a function of weight group and probability that the subject was below his set point, are presented in Table 1. Normal weight subjects were classified as probably below their set points if they had ever been overweight. Underweight subjects were classified as probably below their set points if they were 10% or more under the average weight for their height. Overweight subjects were classified as being at their set points if they were 40% or more overweight. The arbitrariness of the cutoffs for underweight and overweight subjects should be emphasized. The cutoffs were chosen partly because they maximize between-group differences. However, it should be noted that the general pattern of Table 1 holds when virtually any other cutoffs are used. It may be seen that among subjects probably at or near their set points, there is very little difference in eating behavior as a function of weight group. Within each weight group, however, the subjects who are probably below their set points are more responsive to the taste of the ice cream than subjects probably at their set points. In each weight group, subjects below their set points eat more good ice cream and less bad ice cream than their counterparts who are presumably closer to their set points (F Taste × Relation to Set Point interaction $= 7.77$, $p < .01$).

Thus, when an obese individual's degree of overweight is so great as to indicate that he has ignored social pressure and allowed his weight to equal or exceed its set point, he appears to behave like individuals of normal weight. However, when there is evidence that the individual is below his set point, he appears to behave like an obese individual, even if he is of normal weight or is underweight.

A series of experiments on the taste responsiveness of rats carried out by L. R. Hanson and R. E. Nisbett (unpublished data) gives additional support to the "ponderostat" concept. If overweight humans are more taste responsive because of their obese state, or something associated with it, then the more overweight members of nonhuman species would be expected to be more taste responsive than the less overweight. On the other hand, if it is correct that overweight humans are more taste responsive because they are below their set points, then there is no reason to expect the relatively overweight members of nonhuman species, who would have no reason to be below their set points, to be more taste responsive than their fellows. In order to examine this question, Hanson and Nisbett measured taste responsiveness in Hooded, Holtzman, heavy Charles River strain, and inbred Lewis rats. Taste responsiveness was studied at weaning, at maturity, and in old age. Taste tests included quinine adulteration of dextrose solutions and of chow, cellulose dilution of chow, replacement of chow with high fat content food, and a choice between high fat content food and standard chow. In about half of the tests, a significant association was found between *absolute weight* and taste responsiveness, such that larger rats ate more good food or less bad food or chose to consume proportionally more good tasting food. (Comparisons were always within, never between, strains.) A positive association was never found, however, between any measure of taste responsiveness and *degree of overweight*, whether overweight was defined as weight/length^3, weight of inguinal fat pad, fat pad weight/body weight, fat pad weight/length or length^3, or percentage of fat content of the carcass. On the contrary, weak negative associations were sometimes found, such that the more underweight rats were the more taste responsive.

It thus appears that overweight is not inevitably associated with taste responsiveness. Among rats, who of course would not attempt to control obesity, a positive association between relative overweight and taste responsiveness was not found. On the other hand, it would be expected that rats should respond as do humans if they were to undergo a weight reduction program. In order to test this hypothesis, Nisbett, Hanson, Harris, and Stair (1972) insituted a program of modest weight reduction in rats without employing the severe deprivation regimens of previous experiments. Mayer, Marshall, Vitale, Christensen, Mashayekhi, and Stare (1954) have shown that when a laboratory rat is given an hour's worth of exercise each day, it does not maintain its weight. Instead, a small drop in weight occurs at the beginning of the exercise regimen, and the animal eats only enough to sustain itself at its new weight level, rather than eating enough to compensate for the energy loss due to exercise. Why this occurs is a mystery, but it is an extremely useful phenomenon for the purpose of examining the behavior of animals whose weight has temporarily dropped below its previously established level.

Nisbett et al. (1972) subjected animals to an exercise regimen similar to that employed by Mayer et al. (1954), requiring animals to move at a brisk walk in motor driven activity wheels for an hour each day. Taste responsiveness of exercised animals was compared with that of unexercised control animals. Compared to controls, exercised animals exhibited greater saccharin preference, chose to consume proportionally more good tasting, high fat content food than standard chow, increased their caloric intake more from chow base-line levels when offered the high fat food, and decreased their intake from chow base-line levels more when offered only quinine-adulterated chow. These findings seem particularly striking in view of the fact that weight loss due to exercise never exceeded five per cent of original body weight.

The implications of these data for present concerns seem considerable. Weight loss was produced without resort to deprivation schedules and a very minor drop from previously established set point resulted in marked changes in taste responsiveness. It thus appears that the central nervous system is extremely vigilant, responding even to modest incursions on the fat store by pronounced alterations in behavior.

VENTROMEDIAL HYPOTHALAMUS: THE PONDEROSTAT?

It will now be profitable to discuss an issue that will be uppermost in the minds of readers familiar with the behavior of the rat with lesions of the ventromedial hypothalamus. Virtually every aspect of the behavioral parallels between obese and hungry humans seems also to have a parallel to the behavior of the rat with ventromedial hypothalamus lesions. The parallels between the obese human and the ventromedially lesioned rat have been described by Nisbett

(1968b, in press; Nisbett & Gurwitz, 1970) and Schachter (1971), so it is only necessary to briefly sketch the parallels here. No attempt will be made to document extensively the generalizations about the behavior of the ventromedially lesioned rat, since they are well known. The classic references on the eating behavior of the ventromedially lesioned rat include work by Miller, Bailey, and Stevenson (1950), and Teitelbaum (1955, 1957).

Like the obese human, and like the hungry organism, the ventromedially lesioned rat eats more at a given meal, eats more rapidly, and has a greater readiness to eat as evidenced by the fact that it is more likely to eat in a novel environment containing food (Sclafani, 1971; Sclafani, Belluzzi, & Grossman, 1970). Like the obese human, and like the hungry organism, the ventromedially lesioned rat is highly taste responsive. Compared to the normal animal, it dramatically increases its intake over baseline, standard chow levels when offered particularly good-tasting food and decreases its intake below chow levels when offered only bad-tasting food.

It is fair to say that the chief effects of ventromedial hypothalamus lesions on behavior other than eating behavior fall perfectly within the by-now-familiar categories of emotionality, activity, and sex. The ventromedially lesioned animal has long been known for its irritability and proneness to rage reactions in response to the minor routines of daily laboratory handling. Recent evidence indicates that ventromedially lesioned animals are also highly fearful. They learn to avoid electric shock more quickly than normal animals, and their thresholds for flinching, jumping, and vocalizing in response to shock are lower (Grossman, 1966; Sechzer, Turner, and Liebelt, 1966; Turner, Sechzer, & Liebelt, 1967).

Ventromedially lesioned rats have also long been known to be highly inactive. They move around far less both in the home cage and in running wheels. This of course resembles the pattern in human obesity and human hunger but differs from the pattern of the hungry rat. The discrepancy is not surprising, however, in view of the fact that, unlike the deprived rat, the ventromedially lesioned rat has no problem with temperature regulation due to dwindling fat store energy reserves.

Finally, it has long been known that the ventromedially lesioned rat is hyposexual. Ventromedial hypothalamus hyposexuality, like the hyposexuality of hunger, is probably mediated at least in part by a lowering of gonadal hormone levels (see Valenstein, Cox, & Kakolewski, 1969, for a review of this literature). Estrous cycle in the female is frequently disrupted or abolished, and the testicles of males are frequently found in the cryptorchid position. Gonadal weights are typically lower in both sexes following ventromedial hypothalamus lesions. Whether these effects are direct consequences of ventromedial hypothalamus lesions or whether they are the result of damage to neighboring structures is not clearly established.

The triple analogy is thus complete: In major and highly specific ways the behavior patterns of the ventromedially lesioned rat, the obese human, and the food-deprived organism closely resemble one another. It should be noted, however, that it is customary in the eating behavior literature to *deny* that the ventromedially lesioned rat is hungry or behaves as if it were hungry. It is important to discuss the reasons for this.

Historically, the view that the ventromedially lesioned rat is hungry, or rather is confused by the lesion so as to behave as if it is hungry, is the oldest. The fact that the ventromedially lesioned rat ate prodigious amounts of food and became fat was considered to be presumptive evidence in favor of this view. The view was sharply disputed, however, by Miller et al. (1950) on the basis of two observations: the ventromedially lesioned rat is highly taste responsive and it is less willing to work for its food than normal animals.

The taste responsiveness of the ventromedially lesioned rat, of course, cannot today be considered as evidence that it is not hungry. In view of recent research, it may now be taken as evidence that the ventromedially lesioned rat *is* hungry.

A secondary point concerning the finickiness of lesioned animals deserves more detailed consideration, however. The view widely prevails that the "dynamic," not-yet-obese lesioned animal is not highly taste responsive and that finickiness is observed chiefly in the obese, "static" animal. This view is inconsistent with the present position, which holds that hunger, and therefore finickiness, should be at least as great prior to attainment of obesity as after attainment. The view that the dynamic animal is not highly finicky, however, seems to rest primarily on a single study by Teitelbaum (1955). Teitelbaum reported that his dynamic animals ate almost heedlessly, paying much less attention to taste than static animals and only slightly more attention to taste than controls. Teitelbaum suggested that obesity was a precondition for manifestation of maximal finickiness. This view is sharply contradicted by several studies. Very dramatic finickiness has been demonstrated in dynamic animals by Stellar and his colleagues. In one study (Corbit & Stellar, 1964), dynamic animals were offered pellets after weight had stabilized on less-preferred powder. The increases in food intake were far greater for lesioned animals than for control animals, and far greater than those observed by Teitelbaum in a similar test with dynamic animals. Moreover, intake of the preferred pellets *declined* for the dynamic animals as their weight *increased*. A comparable picture was obtained when animals were then switched to a greatly preferred high fat diet. Lesioned animals manifested huge increases in intake which again tapered off as weight increased. A similar decline in taste responsiveness accompanying weight gain has been observed by others for sweet tastes. Both Soulairac (1947) and Jacobs and Sharma (1969) have reported that exaggerated glucose preferences manifested by animals in the dynamic stage tend to disappear in the static phase. The bulk of the data therefore does not support the view that obesity is a precondition for taste responsiveness. This view, in fact, is explicitly contradicted by Stellar and his colleagues:

> Obesity is not a necessary condition for finickiness. Some of our most obese animals were only minimally finicky, and maximal finickiness was seen in animals that were not obese during testing . . . [Graff & Stellar, 1962, p. 424].

The reader should be reminded that the interrelationships among obesity, taste responsiveness, and hyperphagia are very complicated and still poorly understood. (See Mrosovsky, 1971, for an excellent and well-balanced review of the pitfalls and complexities in this area.) Space permits only the above brief summary of the evidence touching on the crucial point of finickiness in the dynamic animal. The available data are probably insufficient to determine whether it is the dynamic or the obese rat which is the more taste responsive. But there is little doubt that dramatic taste responsiveness does occur in the dynamic stage.[3]

The second observation which implied to Miller et al. (1950) that the ventromedially lesioned animal was not hungry was their finding that the animal was unwilling to work for its food. Miller et al. (1950) showed that the ventromedially lesioned animal pulled less hard on a restraining harness to reach food and that it was less willing to push a lever to obtain pellets on a schedule that provided infrequent reinforcement. This is of course not at all the behavior which would be expected of a hungry animal. On the other hand, this behavior should be viewed in context with other facts about the ventromedially lesioned animal. It is highly emotional and highly inactive. The harness tests were performed after 72 hours of total food deprivation. This means that the "control" animals would have been extremely hungry, and probably maximally willing to exert effort to obtain food. The extreme emotionality of the ventromedially lesioned rats, however, as Grossman (1966) pointed out, might have resulted in sufficient frusjects, taking advantage of the fact that, for most Americans, life has provided pretraining in the necessity to unwrap Hershey kisses before eating them. Unwrapping them, then, should not be an arbitrary, unpleasant requirement, but a familiar, comfortable part of the ritual of consuming them. And in fact they found that, with the chocolates, the necessity to unwrap did not lower the number eaten by obese subjects (or by normal subjects). In another experiment, Singh (in press) was able to demonstrate the same phenomenon in an experiment where pretraining was actually manipulated. He found that if obese subjects were pretrained to perform a set of complicated physical actions, they were willing to perform them to obtain food, whereas they were much less willing to do so if they had not been pretrained. Pretraining had virtually no effect on the behavior of normal subjects.

---

[3] The present framework does not require that the static animal be *less* finicky than the dynamic and in fact has little to say about the complicated static preparation which may or may not have some recovery of function and which, though perhaps at its set point so far as its VMH is concerned, may be above it with respect to other central nervous system structures. The present view requires only that the dynamic animal be more finicky than the unoperated animal.

It thus appears that, for obese humans as for ventromedially lesioned rats, the lack of willingness to work for food may not be a reflection of lack of hunger motivation but of decreased frustration tolerance.

A third difficulty for the present framework is presented by the contention of Nisbett and Gurwitz (1970) that the eating behavior of overweight human newborns resembles that of the animal with ventromedial hypothalamus lesions. From the present viewpoint, such a finding would be quite troublesome, since there is no reason to assume that the overweight newborn would be below its set point. In fact, if anything, the opposite should be the case. However, the actual data presented by Nisbett and Gurwitz allow a more congenial interpretation. One of the findings was that overweight infants (infants with high weight/length ^{3} ratios) were unwilling to work for their food, as evidenced by the fact that their intake was suppressed relatively more when they were required to suck through a nipple with a very small hole. This finding of course is more or less consistent with the present framework. It might be expected that the relatively underweight infants would be hungrier and relatively more willing to work for food than the overweight infants.

The other finding reported by Nisbett and Gurwitz (1970) is somewhat more difficult to reconcile. Heavy infants (infants whose absolute weight was greater) were found to be more taste responsive, as evidenced by the fact that they increased their intake relatively more when offered a sweet-tasting formula. However, overweight (weight/length ^{3}) was unrelated to taste responsiveness. Nisbett and Gurwitz attributed this lack of association to the fact that, in the taste study, infant length measurements were unreliable because they were obtained by harried hospital staff who tend to be rather casual in their treatment of such low priority data. It would be more consistent with the present framework to assume that unreliability was not the cause of the lack of association and that the data should be accepted at face value. (The data are quite consistent, incidentally, with the work of Hanson and Nisbett on rats. Absolute weight was sometimes found to be positively related to taste responsiveness, but overweight was not found to be associated with taste responsiveness.)

A follow-up study by Nisbett and Gurwitz (unpublished data) provides some useful information. The study was essentially a replication of both of the procedures of the earlier studies. On two separate days, infants were given three experimental feedings: one standard formula, one sweetened formula, and one hard-to-suck highly viscous formula (Nutramigen, by Mead-Johnson). As in the previous study, overweight was associated with unwillingness to work for food. Infants with high weight/length ^{3} ratios consumed proportionately less hard-to-suck than standard formula (82% as much) compared to infants with medium ratos (87% as much) or small ratios (93% as much). The differences fall short of statistical significance, but the trend is similar to that observed in the original study. And, as in the original study, overweight was unrelated to taste responsiveness. Unfortunately, absolute weight was also found not to be systematically related to taste responsiveness, so it cannot be said that all inconsistencies are resolved by the second study. Nevertheless, the two studies are fairly consistent with respect to the more important independent variable of overweight: in both studies, relative overweight was associated with relative unwillingness to work for food; and in neither study was relative overweight associated with greater taste responsiveness.

Another aspect of the follow-up study is quite pertinent to present concerns. Infants were reweighed one month after birth, and it is therefore possible to examine the relationship between proportional weight gain in the weeks after birth and the patterns of eating behavior at birth. This analysis has relevance for the set point notion if the assumption is made that nutritional conditions may differ in utero such that some infants are relatively deprived at birth. If so, the relatively deprived infants would be expected to gain weight more rapidly when given the opportunity to control their own nutrition. The infants who subsequently gain more might therefore be expected to be quite willing to work for food at birth and to be highly responsive to the taste of food. At any rate, this was the case. High-gain infants were significantly more willing to consume hard-to-suck formula than low-gain infants and increased their intake of sweetened formula significantly more.

A WARNING ABOUT PREMATURE APPLICATION

On balance, the available data do not appear to present problems for the theoretical structure serious enough to warrant more than a cautionary note. However, more than a modicum of caution is recommended with respect to the chief practical implications of the present view. The history of medical opinion on the causes of obesity has been one of cycling, every generation or so, between the psychological "gluttony" explanation and the totally biological, "glandular" extreme. It would be very unfortunate if the effect of this review were to further an uncritical acceptance of the inevitability of obesity or to encourage clinical workers to counsel their obese patients not to attempt to lose weight. This is true for two reasons: (*a*) The present view that the obese may be below their set points has received no direct test as of the present time, and is in fact the weakest link in the present theoretical position. (*b*) An alternative view of the obese, not involving the set point notion, appears at least somewhat tenable.

This alternative view, having to do with the meal frequency pattern of the obese, should be briefly sketched. Cohn has demonstrated that when rats and chickens are fed meals at infrequent intervals, rather than being fed ad lib, they become obese. (See Cohn and Joseph, 1960, for a review.) The work indicates that when calories are consumed in a brief period of time, they are stored as fat, whereas the same calories, when taken in small, frequent meals, do not result in fat buildup. This fact may have relevance for human obesity. Obese humans have been repeatedly found, in eating behavior surveys by Schachter and by Nisbett, to eat infrequently. The obese individual consumes his calories in fewer meals, spaced further apart, than does the normal individual. If Cohn's data are applicable to humans, it could be that some individuals are obese in part because of their habit of concentrated calorie intake. Obese subjects might therefore show hungry patterns of behavior in experimental situations merely because of short-term energy deficit and not at all because of any long-term deficit. Such a view would be inconsistent with a good deal of the data reported throughout this paper, but on the other hand its implications have not been as fully explored as those of the present view. Because the conclusion suggested by the Cohn and Joseph data is that obesity may often represent an elevation of fat stores over set point, and because adequate experimentation might conceivably support this implication, any application of the present theoretical view to clinical practice would be quite inappropriate.

SUMMARY

The starting point for the present inquiry was a set of speculative thoughts prompted by the apparent fact that obesity reflects a biologically programmed elevation of fat store levels above average values. The result of these speculations has been a reevaluation not only of obesity in humans but of the condition of hunger and the

functions of the ventromedial hypothalamus. This integrative effort may be summarized as a set of propositions and subpropositions more or less deductively related to one another.

1. Hunger exerts its effects in great part through an inactivation of the ventromedial hypothalamus, since: (*a*) Food deprivation results in morphological evidence of physiological inactivity of the ventromedial hypothalamus. (*b*) Eating is followed by electrical activity of the ventromedial hypothalamus, and food deprivation by electrical inactivity of the ventromedial hypothalamus. (*c*) Lesion of the ventromedial hypothalamus produces essentially an exaggerated version of the behavioral effects of hunger. (*d*) Stimulation of the ventromedial hypothalamus causes cessation of eating.

2. The obese individual displays the behavior characteristic of hunger and of ventromedial hypothalamus lesions, suggesting either that the obese individual is hungry or has a weak ventromedial hypothalamus.

3. The evidence favors the possibility that the obese individual is hungry because: (*a*) Free fatty acid levels are elevated in naturally occurring obesity as in hunger. (*b*) Free fatty acid levels are decreased in artificially induced obesity. (*c*) The very fattest individuals show little evidence of the "hungry" syndrome. (*d*) The moderately obese individual and the formerly obese individual show the greatest evidence of the "hungry" syndrome.

The most salient theoretical advantage of the present view is that, while doing little violence to the existing data, it provides common theoretical constructs with which to integrate phenomena in three disparate areas of eating behavior research: hunger, obesity, and hypothalamic hyperphagia.

REFERENCES

Anand, G. K., Chhina, G. S., & Singh, B. Effect of glucose on the activity of hypothalamic "feedings centers." *Science*, 1962, **138**, 597–598.

Anand, B. K., Chhina, G. S., Sharma, K. N., Dua, S., & Singh, B. Activity of single neurons in the hypothalamic feeding centers: Effect of glucose. *American Journal of Physiology*, 1964, **207**, 1146–1154.

Bacon, W. E., Snyder, H. L., & Hulse, S. H. Saccharin preference in satiated and deprived rats. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 112–114.

Bagdade, J. D., Porte, D., & Bierman, E. L. The interaction of diabetes and obesity on the regulation of fat mobilization in man. *Diabetes*, 1969, **18**, 759–772.

Björntorp, P. Disturbances in the regulation of food intake. *Advances in Psychosomatic Medicine*, 1972, in press.

Björntorp, P., Bergman, H., & Varnauskas, E. Plasma free fatty acid turnover rate in obesity. *Acta Medica Scandanavia*, 1969, **185**, 351–356.

Björntorp, P., Bergman, H., Varnauskas, E., & Lindholm, B. Lipid mobilization in relation to body composition in man. *Metabolism*, 1969, **18**, 840–851.

Brown, J., D., & Pulsifer, D. H. Outpatient starvation in normal and obese subjects. *Aerospace Medicine*, 1965, March, 267–269.

Brown, K. A., & Melzack, R. Effects of glucose on multi-unit activity in the hypothalamus. *Experimental Neurology*, 1969, **24**, 363–373.

Bruch, H. *The importance of overweight.* New York: Norton, 1957.

Bullen, B. A., Reed, R. B., & Mayer, J. Physical activity of obese and nonobese adolescent girls appraised by motion picture sampling. *American Journal of Clinical Nutrition*, 1964, **14**, 211–223.

Cabanac, M., & Duclaux, R. Obesity: Absence of satiety aversion to sucrose. *Science*, 1970, **168**, 496–497.

Cabanac, M., Duclaux, R., & Spector, N. H. Sensory feedback regulation of body weight: Is there a ponderostat? *Nature*, 1971, **229**, 125–127.

Chirico, A., & Stunkard, A. J. Physical activity and human obesity. *New England Journal of Medicine*, 1960, **263**, 935–940.

Cohn, C., & Joseph, D. Effects on metabolism produced by the rate of ingestion of the diet. *American Journal of Clinical Nutrition*, 1960, **8**, 682–689.

Corbit, J. D., & Stellar, E. Palatability, food intake and obesity in normal and hyperphagic rats. *Journal of Comparative and Physiological Psychology*, 1964, **58**, 63–67.

Corvilain, J., Loeb, H., Champenois, A., & Abramow, M. Effect of fasting on levels of plasma-nonesterified fatty acids in normal children, normal adults, and obese adults. *Lancet*, 1961, March 11, 534–535.

Davenport, C. B. Body-build and its inheritance. (Publication No. 329.) Washington, D. C.: Carnegie Institution of Washington, 1923.

Dole, V. P. A relation between nonesterified fatty acids in plasma and the metabolism of glucose. *Journal of Clinical Investigation*, 1956, **35**, 150–154.

Fuller, J. L., & Jacoby, G. A. Central and sensory control of food intake in genetically obese mice. *American Journal of Physiology*, 1955, **183**, 279–283.

Genuth, S. M. Effects of prolonged fasting on insulin secretions. *Diabetes,* 1966, **15**, 798–806.

Gliddon, J. B. Effect of glucose, glucose-6-phosphate, and fructose-6-phosphate on the electrical activity in the ventromedial hypothalamus in rats. *Brain Research,* 1970, **22**, 429–433.

Gluckman, M. L., & Hirsch, J. The response of obese patients to weight reduction. *Psychosomatic Medicine,* 1968, **30**, 1–11.

Gluckman, M. L., Hirsch, J., McGully, R. S., Barron, B. A., & Knittle, J. L. The response of obese patients to weight reduction. *Psychosomatic Medicine,* 1968, **30**, 359–373.

Gollnick, P. D. Exercise, adrenergic blockage and free fatty acid mobilization. *American Journal of Physiology,* 1967, **213**, 734–738.

Gordon, E. S. Nonesterified fatty acids in the blood of obese and lean subjects. *American Journal of Clinical Nutrition,* 1960, **8**, 740–747.

Gordon, E. S. Metabolic aspects of obesity. In R. Levine & R. Luft (Eds.), *Advances in metabolic disorders.* New York: Academic Press, 1970.

Gordon, R. S., & Cherkes, A. Unesterified fatty acids in human blood plasma. *Journal of Clinical Investigation,* 1956, **35**, 206–212.

Goss, R. J. Hypertrophy versus hyperplasia. *Science,* 1966, **153**, 1516–1620.

Graff, H., & Steller, E. Hyperphagia, obesity and finickiness. *Journal of Comparative and Physiological Psychology,* 1962, **55**, 418–424.

Gross, L. P. Scarcity, unpredictability and eating behavior in rats. Unpublished doctoral dissertation, Columbia University, 1968.

Grossman, S. P. The VMH: A center for affective reactions, satiety, or both? *Physiology and Behavior,* 1966, **1**, 1–10.

Hamilton, C. L. Problems of refeeding after starvation in the rat. *Annals of the New York Academy of Sciences,* 1969, **157**, 1004–1017.

Hamilton, C. L., & Brobeck, J. R. Hypothalamic hyperphagia in the monkey. *Journal of Comparative and Physiological Psychology,* 1964, **57**, 271–278.

Hanson, P. G., Johnson, R. E., & Zaharto, D. S. Correlation between ketone body and free fatty acid concentrations in the plasma during early starvation in man. *Metabolism,* 1965, **14**, 1037–1040.

Hashim, S., & Van Itallie, T. Studies in normal and obese subjects with a monitored food-dispensing device. *Annals of the New York Academy of Sciences,* 1965, **131**, 654–661.

Heald, F. P., Mueller, P. S., & Daugela, M. Z. Glucose and free fatty acid metabolism in obese adolescents. *American Journal of Clinical Nutrition,* 1965, **16**, 256–264.

Hervey, G. R. Regulation of energy balance. *Nature,* 1969, **222**, 629–631.

Hirsch, J., & Han, P. W. Cellularity of rat adipose tissue: Effects of growth, starvation, and obesity. *Journal of Lipid Research,* 1969, **10**, 77–82.

Hirsch, J., & Knittle, J. L. Cellularity of obese and nonobese human adipose tissue. *Federation Proceedings,* 1970, **29**, 1516–1521.

Hoebel, B. G., & Teitelbaum, P. Weight regulation in normal and hypothalamic hyperphagic rats. *Journal of Comparative and Physiological Psychology,* 1966, **61**, 189–193.

Hoff, L. A., & Gentile, R. L. Deprivation and the role of taste in food intake. *Psychonomic Science,* 1969, **17**, 277–278.

Issekutz, B., Bortz, W. M., Miller, H. I., & Wroldsen, A. Plasma free fatty acid response to exercise in obese humans. *Metabolism,* 1967, **16**, 492–502.

Jacobs, H. L., & Sharma, K. N. Taste versus calories: Sensory and metabolic signals in the control of food intake. *Annals of the New York Academy of Sciences,* 1969, **157**, 1084–1125.

Jahnke, K., Gries, F. A., Wallenfels, H., & Schulte, H. Verhalten von Metaboliten des Fettstoffwecksels im Serum adipöser und nicht-adipöser Personen unter Grundumsatzbedingungen. *Klinische Wochenschrift,* 1964, **42**, 1016–1020.

Kennedy, G. C. The development with age of hypothalamic restraint upon the appetite of the rat. *Journal of Endocrinology,* 1957, **16**, 9–17.

Keys, A., Brožek, J., Henschel, A., Mickelsen, O., & Taylor, H. L. *The biology of human starvation.* Minneapolis: University of Minnesota Press, 1950.

Knittle, J. L., & Hirsch, J. Effect of early nutrition on the development of rat epididymal fat pads: Cellularity and metabolism. *Journal of Clinical Investigation,* 1968, **47**, 2091.

Larkin, R. P. How hard will hyperphagic rats logical Association meetings, New York, April 1971.

Leyton, G. B. Effects of slow starvation. *Lancet,* 1946, July 20, 73–78.

Liebermeister, H., Daweke, H., Gries, F. A., Schilling, W. H., Grüneklee, D., Probst, G., & Jahnke, K. Einflüss der Gewichtsreduktion auf Metabolite des Kohlenhydrat- und Fettstoffwechesels und auf das Verhalten des Seruminsulins bei Adipösitas. *Diabetologia,* 1968, **4**, 123–132.

Mager, M., & Iampietro, P. F. The effect of prolonged cold and starvation and subsequent refeeding on plasma lipids and glucose of normal men. *Metabolism,* 1966, **15**, 9–16.

Marshall, M. W., Smith, B. P., Munson, A. W., & Lehmann, R. P. Prediction of carcass fat from body measurements made on live rats differing in age, sex and strain. *British Journal of Nutrition,* 1969, **23**, 353–369.

Mayer, J. Regulation of energy intake and the body weight: The glucostatic theory and the lipostatic hypothesis. *Annals of the New York Academy of Science,* 1955, **63**, 15–43.

Mayer, J., Marshall, N. B., Vitale, J. J., Christensen, J. H., Mashayekhi, M. B., & Stare, F. J. Exercise, food intake and body weight in normal rats and genetically obese adult mice. *American Journal of Physiology*, 1954, **177**, 544–548.

MEERLOO, J. A. M., & KLAUBER, L. D. Clinical significance of starvation and oral deprivation. *Psychosomatic Medicine*, 1952, **14**, 491–497.

MILLER, E. V. O., MICKELSEN, O., & KEYS, A. Urinary excretion of 17-ketosteroids by normal young men during starvation. *Proceedings of the Society for Experimental Biology and Medicine*, 1948, **67**, 288–292.

MILLER, N. E., BAILEY, C. J., & STEVENSON, J. A. F. Decreased "hunger" but increased food intake resulting from hypothalamic lesions. *Science*, 1950, **112**, 256–259.

MOORE, M. E., STUNKARD, A., & SROLE, L. Obesity, social class, and mental illness. *Journal of the American Medical Association*, 1962, September, 138–142.

MORSE, W. I., & MAHABIR, R. Changes in glucose tolerance and plasma free fatty acids after fasting in obesity. *Diabetes*, 1964, **13**, 286–290.

MROSOVSKY, N. *Hibernation and the hypothalamus.* New York: Appleton-Century-Crofts, 1971.

NISBETT, R. E. Taste, deprivation, and weight determinants of eating behavior. *Journal of Personality and Social Psychology*, 1968, **10**, 107–116. (a)

NISBETT, R. E. Determinants of food intake in human obesity. *Science*, 1968, **159**, 1254–1255. (b)

NISBETT, R. E. Eating behavior and obesity in men and animals. *Advances in Psychosomatic Medicine*, in press.

NISBETT, R. E., & GURWITZ, S. Weight, sex, and the eating behavior of human newborns. *Journal of Comparative and Physiological Psychology*, 1970, **73**, 245–253.

NISBETT, R. E., HANSON, L. R., HARRIS, A., & STAIR, A. Taste responsiveness, weight loss and the ponderostat. Unpublished manuscript, University of Michigan, 1972.

NISBETT, R. E., & KANOUSE, D. E. Obesity, food deprivation, and supermarket shopping behavior. *Journal of Personality and Social Psychology*, 1969, **12**, 289–294.

OOMURA, Y., OOYAMA, H., NAKA, F., YAMAMOTO, T., ONO, T., & KOBAYASHI, N. Some stochastical patterns of single unit discharges in the cat hypothalamus under chronic conditions. *Annals of the New York Academy of Science*, 1969, **157**, 666–689.

OOMURA, Y., OOYAMA, H., YAMAMOTO, T., & NAKA, F. Reciprocal relationship of the lateral and ventromedial hypothalamus in the regulation of food intake. *Physiology and Behavior*, 1967, **2**, 97–115.

OPIE, L. H., & WALFISH, P. G. Plasma free fatty acid concentrations in obesity. *New England Journal of Medicine*, 1963, **268**, 757–760.

PFAFF, D. W. Histological differences between ventromedial hypothalamic neurones of well fed and underfed rats. *Nature*, 1969, **223**, 77–78.

POWLEY, T. L. Hypothalamic feeding centers control adipose tissue mass. Paper presented at the Fourth International Conference of the Regulation of Food and Water Intake, Cambridge, England, August 1971.

POWLEY, T. L., & KEESEY, R. Relationships of body weight to the lateral hypothalamic feeding syndrome. *Journal of Comparative and Physiological Psychology*, 1970, **70**, 25–36.

RABINOWITZ, D., & ZIERLER, K. L. Forearm metabolism in obesity and its response to intraarterial insulin. *Journal of Clinical Investigation*, 1962, **41**, 2173–2181.

RODIN, J. Shock avoidance behavior in obese and normal subjects. Unpublished manuscript, Yale University, 1972.

RODIN, J., ELMAN, D., & SCHACHTER, S. Emotionality and obesity. Unpublished manuscript, Yale University, 1972.

ROSS, L. D. Cue and cognition controlled eating among obese and normal subjects. Unpublished doctoral dissertation, Columbia University, 1969.

RYTAND, D. A. Hereditary obesity of yellow mice: A method for the study of obesity. *Proceedings of the Society for Experimental Biology and Medicine*, 1943, **54**, 340–341.

SCHACHTER, S. Cognitive labeling of bodily state: Studies of hunger and obesity. In D. Glass (Ed.), *The neurophysiology of emotion.* New York: The Russell Sage Foundation and the Rockefeller University Press, 1967.

SCHACHTER, S. Obesity and eating. *Science*, 1968, **161**, 751–756.

SCHACHTER, S. Some extraordinary facts about obese humans and rats. *American Psychologist*, 1971, **26**, 129–144.

SCHACHTER, S., GOLDMAN, R., & GORDON, A. Effects of fear, food deprivation, and obesity on eating. *Journal of Personality and Social Psychology*, 1968, **10**, 107–116.

SCHEMMEL, R., MICKELSEN, O., & GILL, J. L. Dietary obesity in rats: Influence of diet, weight, fat accretion in seven strains of rats. *Journal of Nutrition*, 1970, **100**, 1041–1048.

SCHEMMEL, R., MICKELSEN, O., & TOLGAY, C. Dietary obesity in rats: influence of diet, weight, age, and sex on body composition. *American Journal of Physiology*, 1969, **216**, 373–379.

SCLAFANI, A. Behavioral effects of selective destruction of ventromedial hypothalamic neuropathways in the rat. Unpublished manuscript, Brooklyn College, 1971.

SCLAFANI, A., BELLUZZI, J. D., & GROSSMAN, S. P. Effects of lesions in the hypothalamus and amygdala on feeding behavior in the rat. *Journal of Comparative and Physiological Psychology*, 1970, **72**, 394–403.

SECHZER, J. A., TURNER, S. G., & LIEBELT, R. A. Motivation and learning in mice after goldthioglucose-induced hypothalamic lesions. *Psychonomic Science*, 1966, **4**, 259–260.

SIMS, E. A., KELLEHER, P. E., HORTON, E. S., GLUCK, C. M., GOODMAN, R. F. & ROWE, D. A. Experimental obesity in man. *Excerpta Medica Monograph*, 1968.

SINGH, D. Role of response habits and cognitive factors in determination of behavior of obese humans. *Journal of Personality and Social Psychology*, in press.

SINGH, D. Hunger motivation in hypothalamic hyperphagic rats. Unpublished manuscript, University of Texas, 1972.

SINGH, D., & SIKES, S. Hunger motivation in obese human: Role of previous experience. Unpublished manuscript, University of Texas, 1972.

SOULAIRAC, A. La physiologie d'un comportement: L'appetit glucidique et la régulation neuro-endocrinienne chez les Rongeurs. *Bulletin de Biologie*, 1947, **81**, 273–432.

SROLE, L., LANGNER, T. S., MICHAEL, S. T., OPLER, M. K., & RENNIE, T. A. C. *Mental health in the metropolis: The midtown Manhattan study*. New York: McGraw-Hill, 1962.

STEFFENS, A. B. Bloodglucose and FFA levels in relation to the meal pattern in the normal rat and the ventromedial hypothalamic lesioned rat. *Physiology and Behavior*, 1969, **4**, 215–225.

STEVENSON, J. A. F., & RIXON, R. H. Environmental temperature and deprivation of food and water on the spontaneous activity of rats. *Yale Journal of Biology*, 1957, **29**, 575–584.

STUNKARD, A. J. The "dieting depression." *American Journal of Medicine*, 1957, July, 77–86.

TEITELBAUM, P. Sensory control of hypothalamic hyperphagia. *Journal of Comparative and Physiological Psychology*, 1955, **48**, 156–163.

TEITELBAUM, P. Random and food-directed activity in hyperphagic and normal rats. *Journal of Comparative and Physiological Psychological Psychology*, 1957, **50**, 486–490.

TENEN, S. S., & MILLER, N. E. Strength of electrical stimulation of lateral hypothalamus, food deprivation, and tolerance for quinine in food. *Journal of Comparative and Physiological Psychology*, 1964, **58**, 55–62.

TURNER, S. G., SECHZER, J. A., & LIEBELT, R. A. Sensitivity to electric shock after ventromedial hypothalamic lesions. *Experimental Neurology*, 1967, **19**, 236–244.

VALENSTEIN, E. S. Selection of nutritive and nonnutritive solutions under different conditions of need. *Journal of Comparative and Physiological Psychology*, 1967, **63**, 429–433.

VALENSTEIN, E. S., COX, V. C., & KAKOLEWSKI, J. W. Sex differences in hyperphagia and body weight following hypothalamic damage. *Annals of the New York Academy of Science*, 1969, **157**, 1030–1048.

VAN ITALLIE, T. B., & HASHIM, S. A. Biochemical concomitants of hunger and satiety in man. *American Journal of Clinical Nutrition*, 1960, **8**, 587–594.

WALKER, D. W., & REMLEY, N. R. The relationships among percentage body weight loss, circulating free fatty acids and consummatory behavior in rats. *Physiology and Behavior*, 1970, **5**, 301–309.

WIDDOWSON, E. M., & MCCANCE, R. A. Some effects of accelerating growth. I. General somatic development. *Proceedings of the Royal Society*, 1960, **152**, 188–206.

WILLIAMS, R. A., & CAMPBELL, B. A. Weight loss and quinine-milk ingestion as measures of "hunger" in infant and adult rats. *Journal of Comparative and Physiological Psychology*, 1961, **54**, 220–222.

(Received April 11, 1972; revision
received July 10, 1972)

Psychological Review
1972, Vol. 79, No. 6, 454–470

FOUNDATIONS OF AN ATTRIBUTIONAL THEORY OF PERFORMANCE[1]

ANDY KUKLA[2]

University of Toronto

A theory of the determinants of performance intensity is outlined which combines concepts of decision theory with those of attribution theory. The theory is used to analyze the effects on the vigor with which a task is undertaken of a number of interrelated variables: the task's perceived difficulty, the subject's perception of his own ability, and the experience of success or failure. The personality dispositions of need for achievement and anxiety are also related to these basic concepts. The relationship between the present theory and Atkinson's theory of achievement motivation is discussed. Finally, some indications are given of possible further developments of this theoretical approach.

A cognitive theory of the contemporary determinants of task performance is presented below. The performance variable of main interest is that of behavioral *intensity*, that is, the vigor with which a person exerts himself toward his selected goal. The related problem of goal selection, or the *direction* of behavior, is dealt with only peripherally.

The present theory is cognitive in the sense that it ascribes motivational significance to people's systems of beliefs concerning themselves and their environment. More specifically, it is assumed that the way people attribute the causes of task outcomes is an important determinant of the characteristics of their performance on these tasks.

The attributional theory of performance is conceptually related to two influential but hitherto independent traditions of cognitive theorizing: attribution theory and expectancy theory. In essence, the present analysis constitutes a logical extension of expectance theory which incorporates the new but unsystematic insights obtained by attribution theory.

Attribution Theory

The conceptual apparatus of the attributional theory of performance incorporates the analysis of causal attribution offered by Heider (1958). In his "naive analysis of action," Heider outlines the manner in which a relatively precise belief system is used to account for the success or failure of undertaken tasks. Four major attributional variables are isolated by this analysis. A particular task outcome may be attributed to the degree of *ability* possessed by the actor, to the amount of *effort* he expends, to the degree of *difficulty* involved in achieving the outcome, or to uncontrollable *chance* factors in the environment. More generally, attributions can be made to any weighted combination of these four factors.

The large body of research generated by the Heiderian analysis of action has come to be called attribution theory (Kelley, 1967). Most of this work has attempted to delineate the antecedent stimulus conditions which determine specific causal attributions. The attributional theory of performance, however, is primarily concerned with the relationship between causal attribution and subsequent action. It attempts to specify how a given pattern of attribution to ability, effort, difficulty, or chance determines task performance. An earlier attempt to construct such an attributional theory of motivated behavior was made by Weiner, Frieze, Kukla, Reed, Rest, and Rosenbaum (1972). The earlier analysis, however, limited itself to informal derivations from not always explicit attributional assumptions. In contrast, the present theory is developed in a formal manner.

---

[1] This paper was completed while the author was the recipient of Canada Council Research Grant S71-0897.

[2] Requests for reprints should be sent to Andy Kukla, Department of Psychology, Scarborough College, University of Toronto, West Hill, Ontario, Canada.

Expectancy Theory

The attributional theory of performance is also a logical extension of expectancy theory. We take the latter term to refer to a distinctive family of cognitive theories which, despite important differences, find themselves in substantial agreement concerning the nature of the basic concepts required for the analysis of behavior. The theoretical positions of Tolman (1932, 1959), Lewin (1935, 1938), Rotter (1954, 1955), Atkinson (1957, 1964), and the decision theorists (Edwards, 1954, 1961) are included in this family. All expectancy theories have their foundation in a common analysis of choice situations. A person is conceived by them to be confronted with a set of alternatives for action. Each of these alternatives is further thought to be associated with a number of possible consequences. These consequences are in turn assumed to have a characteristic degree of attractiveness or unattractiveness to the actor and also a certain subjective likelihood of occurrence given that a particular action is undertaken. Finally, the likelihoods and attractions of each possible outcome of an act combine to determine a unique numerical property of the act. Predictions about behavior are then based in some way on this last quantified characteristic of alternative actions.

In decision theory, for example, the perceived likelihood of an outcome is known as the *subjective probability* of its occurrence, and its attractiveness as the *utility* of its occurrence. The subjective probabilities and utilities of all possible outcomes of an act X combine to determine the act's *subjectively expected utility* (SEU) according to the following equation:

$$SEU_X = P_1U_1 + P_2U_2 + \cdots + P_nU_n,$$

where P_i is the subjective probability of the ith possible outcome of X and U_i is the utility of that outcome. Other expectancy theories use different vocabularies to express essentially the same notions.[3]

The attributional theory of performance described below adopts the basic conceptual scheme of expectancy theory, but extends it so that the more complex cognitions discussed by Heider (1958) are also taken into account. Thus, it effects a theoretical unification of the established, relatively rigorous, but conceptually restricted, viewpoint of expectancy theory with the newer, heuristic, but as yet unsystematic, approach of attribution theory.

CHOICE AND INTENSITY PRINCIPLES IN EXPECTANCY THEORY

The attributional theory of performance can be formulated as an extension of any particular expectancy theory, since it does not try to resolve the issues upon which these theories disagree. For the sake of consistency, only the vocabulary of decision theory is used below.

Choice Principles

A *choice principle* is a postulate which specifies the conditions under which a particular action X is selected from among the available alternatives for action. Although a number of different principles of choice have been hypothesized within the framework of decision theory, these principles have typically been elaborations of the following basic rule: an action X_i is selected from the set of alternative actions $X_1, X_2, \cdots, X_n$ with a probability which is an increasing function of the SEU of X_i and a decreasing function of the sum of the SEUs of all the alternatives $X_1, X_2, \cdots, X_n$. Thus, the alternative likeliest to be selected is simply the one with the highest SEU (see Edwards & Tversky, 1967, p. 316, on "constant utility" models).

The present study is only incidentally concerned with choice behavior. The analysis of intensity phenomena which follows is compatible with, but does not entail, a choice principle of the foregoing type.

---

[3] See Feather (1959) for documentation of the claim that expectancy theories are equivalent with respect to their most fundamental concepts.

Intensity Principles

An *intensity principle* specifies the vigor or magnitude of effort with which a previously selected action is undertaken. The decision-theoretical literature has little to say about such principles. No doubt this is because decision theorists have for the most part dealt with a restricted range of experimental situations within which the behavioral parameter of intensity is unimportant. The object of decision-theoretical investigation has usually been the behavior of choice among a set of alternative games of chance. But it is in just such gambling situations that the intensity dimension of behavior can be ignored: one bets on one number or another of a roulette wheel, and there is no question of betting more or less vigorously. To be sure, decision theory also treats the topic of "riskless choice" (e.g., Adams & Fagot, 1959). But here too, the experimental situations analyzed are the kinds which are fully characterized by directional properties alone: one chooses either a cash prize or an electrical appliance, and it makes no sense to ask about the intensity with which the cash prize is accepted.

Yet choices are more generally made among courses of action whose outcomes depend on the actor's degree of skill and expenditure of effort. A person may for example choose to work on a test of mathemathical aptitude. In this case, his performance is not satisfactorily characterized by stating only the direction of his behavior. The score he obtains on the mathematics test is also a problem for psychological theory, in a way that one's outcome at a game of roulette is not. Furthermore, this score will depend at least in part on the intensity with which the task is undertaken. Thus, a general behavior theory must postulate an intensity principle as well as a choice principle if it is to be adequate for the analysis of situations where skill is a determinant of outcome. The theory outlined below aims at specifying such a principle of intensity within the conceptual framework of decision theory.

AN ATTRIBUTIONAL THEORY OF PERFORMANCE

Basic Concepts

The attributional theory of performance begins, as do all expectancy theories, by assuming that a subject is confronted with a number of possible courses of action. Each of these actions may in turn have any number of possible *outcomes* associated with it. The subject is then assumed to construct a *utility function* (U) which associates a utility to each outcome (of any action). The theory is compatible with any decision-theoretical view concerning how one experimentally discovers the structure of utility functions (see Mosteller & Nogee, 1951, for a typical treatment of this problem).

At this point, classical decision theory assumes that the subject also constructs, for each action X, a *subjective probability function* (P_X) which specifies the probability of each possible outcome of X if X is selected. Thus the probability function, like the utility function, is conceived to be a cognition which procedes any decision making: on the basis of the information available concerning the task's structure and his own dispositions and capacities, the subject calculates *SEU*s and then decides on his course of action accordingly. Yet, in situations where skill is relevant to outcome, that is, in just those situations for which an intensity principle is needed, this is a highly unlikely account of the sequence of events leading to action. If a person is confronted with the option of taking a mathematics test, his subjective likelihood of succeeding is determined not only by available cues but also by *how much effort he intends to exert at the task.* A person may well think that success is virtually assured if only he tries as hard as he can to succeed, and that his probability of succeeding is relatively low if he expends only a moderate effort. If this person decides to try as hard as he can, his probability of succeeding (P_s) will be very high; if instead he decides to exert only a moderate amount of effort, his P_s will be low. Consequently, in a skill situation, P_s is not simply a percept which precedes and determines all decision making; rather it is itself the result of a prior decision making process. Specifically, P_s *is in part determined by the intended intensity of behavior.*

Thus, in classical decision theory, it is assumed that the predecisional information which determines behavior is the utility and P_s of each possible outcome of the available actions. In the attributional theory of performance, this predecisional information is hypothesized to be of a more complex character: for any available action X, it consists of the utility of each outcome of X, and a *function* which specifies what the P_s of each outcome is *for each possible level of exertion.* Consequently, the P_s of an outcome will be known to a subject only after he has decided how much effort he will expend on the task.

The above analysis dictates the following formal statement. For each action X, the subject constructs a function F_X which associates to each possible level of *effort* (E) a unique probability function P_{X_E}; that is, $F_X(E) = P_{X_E}(x_i)$. $P_{X_E}(x_i)$ is then the subjective probability of obtaining the outcome x_i to action X, given an effort of magnitude E.

To each level E of exertion on task X, there then corresponds a distinct subjectively expected utility. If there are n possible outcomes to X, SEU is specified by:

$$SEU_{X_E} = \sum_{i=1}^{n} P_{X_E}(x_i) \cdot U(x_i).$$

If the result of X is perceived to be independent of the amount of effort expended, we have $P_{X_{E_i}}(x) = P_{X_{E_j}}(x)$ for all outcomes x of X and for all i, j. It follows that $SEU_{X_{E_i}} = SEU_{X_{E_j}}$ for all i, j. That is, SEU has the same value for all effort levels. In this case, it is appropriate to talk of the SEU of actions without reference to effort, and the present theory reduces to classical decision theory.

Intensity Principle

Assume that an action X has already been selected, according to some here unformulated choice principle. The *intensity* I_X with which X is undertaken is then postulated to be an increasing function of E^*, the degree of effort which the subject decides to expend. Thus, we distinguish here the actual amount of effort expended (I_X) from the amount of effort the subject intends to expend (E^*). This distinction recognizes the fact that intentions are not always realized. In general, however, it will be reasonable to suppose that these two variables are highly correlated.

The decision procedure for obtaining E^* consists of two stages: (*a*) the set of all effort levels E which yield a maximum value for SEU_{X_E} is formed, and (*b*) E^* is then the least E in this set. In short, the subject picks the least effortful way to get the greatest return.

The following is an illustrative derivation from the intensity principle. In the case, discussed in the previous section, where the result of undertaking action X is independent of the amount of effort expended, the SEU of X is identical for all values of E. Thus, (*a*) the set of all effort levels which yield a maximum value for SEU_{X_E} contains all possible effort levels, and (*b*) the least such effort level is then $E^* = 0$. That is, when outcome is perceived to be independent of effort, subjects will decide to expend no effort at all.

Simplifying Assumptions

A number of assumptions are made solely to facilitate the derivation of hypotheses in the later sections. Suppose that the action X has already been selected.

1. It is assumed that to each effort level E there is associated a specific outcome x_E of X with subjective certainty. In reality, each effort level determines a probability distribution of outcomes. Although this restriction is severe, the hypotheses discussed below can as well be derived without it. The derivations would be tedious, however.

2. Assume that the intensity with which X is in fact performed is a positive linear function of the intended effort level, that is, $I_X = kE^*$ $(k > 0)$. Again, while this simplifies subsequent derivations, only the weaker assumption that I_X is an increasing function of E^* is necessary. Recall that

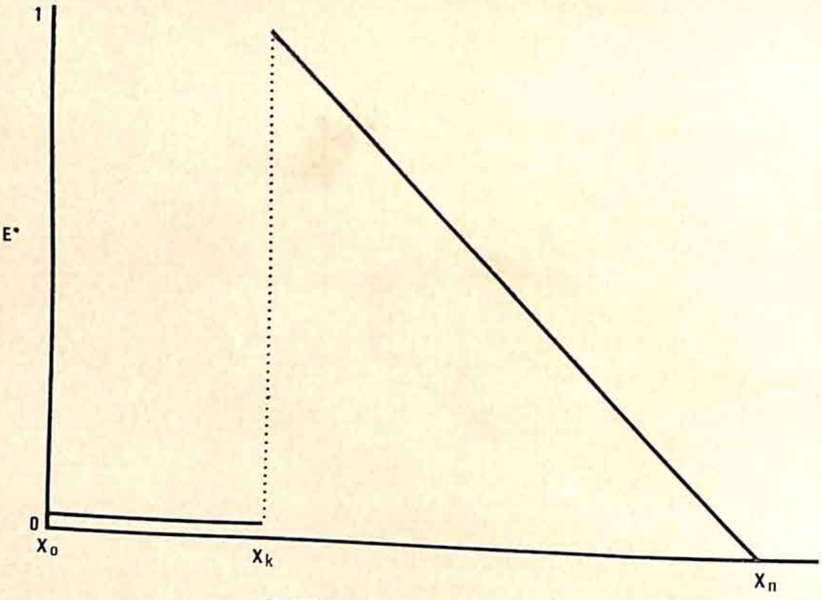


FIG. 2. Theoretical relationship between perceived task facility and intended effort (E^*).

ceived to be more difficult than X_i. The effort–outcome relationship for X_j is a discontinuous step function similar to the one shown in Figure 1. However, by the relationship between effort and perceived difficulty postulated in the previous section, the least effort which results in success at X_j will be greater than that which results in success at X_i. That is to say, the point of discontinuity in the effort–outcome function will be further to the right on the effort scale for X_j than it was for X_i. Again the theory predicts that the selected effort level E_j^* for task X_j will be the point on the effort scale at which the discontinuity occurs. Thus, subjects will decide to expend more effort on X_j than on X_i.

As the tasks considered become more difficult, it is clear that the selected effort levels increase in magnitude. Eventually, the task is perceived to be so difficult that only the maximum possible exertion is expected to result in success. For any task X_k which is still more difficult than this, all possible effort levels are associated with the same failing outcome f_k. In that case, all possible effort levels also result in the same *SEU*. It follows that subjects will select the least such effort level, that is, $E_k^* = 0$. Evidently, the selected effort for any task more difficult than X_k will also be zero.

It is then possible to graph the relationship between perceived task difficulty and intended effort (Figure 2). As long as the task is perceived to be so difficult that not even maximal effort will result in success, the selected effort will be $E^* = 0$. When the task is just easy enough for maximal effort to lead to success, the intended expenditure of effort will be $E^* = 1$. As tasks become easier after this point, E^* goes down (although not necessarily in the linear fashion shown in the figure). For the easiest possible task X_n, success is assured with no effort at all and the subject again does not exert himself.

Since it has been assumed that intensity is a positive linear function of intended effort, Figure 2 depicts the shape of the functional relationship between perceived facility and behavioral intensity as well as that between facility and E^*.

Figure 2 is the intended effort and the intensity function for an individual subject. In a group of subjects, it is to be expected that different individuals will perceive the same task as having varying degrees of facility. Thus, their performance will drop from maximal to minimal intensity (point X_k in Figure 2) at different places along the continuum of tasks. This will result in the group curve's approximating an inverted U rather than the sharply discontinuous function shown in the figure.

Verifying Data

Atkinson (1958, Ch. 10) reports an experiment in which female subjects worked on a skill task for a cash reward of either \$1.25 or \$2.50. Each subject was told how many other persons were competing against her and the number of prizes that would be awarded. One group of subjects was informed that they were in a group of 20 people and that only the best performance of these 20 be rewarded. A second group was instructed that they were competing against two other subjects and that only one of these three could win. Other subjects were told that they had been paired with one other person and that the higher of the two scores would win. Finally, a fourth group was led to believe that they were in a group of four subjects of which the three highest scorers would win a prize. In sum, the proportion of subjects who could succeed at this task varied among the experimental groups from 1/20 through 3/4.

A hypothesis of Heider's suggests a direct relationship between these experimental conditions and the variable of perceived difficulty. According to Heider (1958),

> if we know that practically everyone who tries succeeds, we shall attribute their success to the task. The task is then described as being easy. If hardly anyone succeeds it is felt to be difficult [p. 89].

Frieze and Weiner (1971) have presented evidence corroborating this highly plausible conjecture. Thus, the several groups in Atkinson's (1958) experiment were confronted with tasks whose perceived facility increased as the proportion ot winners increased. The previous analysis of perceived difficulty then leads to the prediction that the intended effort of these groups, and thus also the intensity of their performance, will be related to the proportion of winners in the manner depicted in Figure 2.

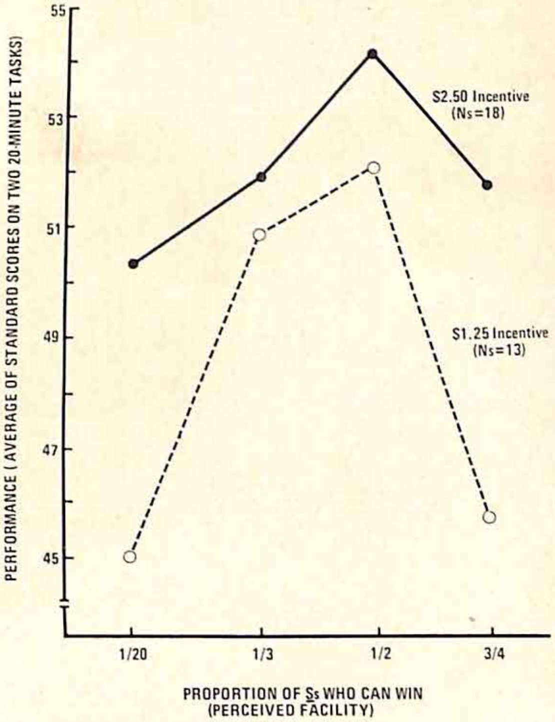


FIG. 3. Observed relationship between intensity of performance and perceived task facility, with monetary incentives of \$1.25 and \$2.50 (adapted from Atkinson, 1958, Ch. 10, p. 294. From Motives in Fantasy, Action, and Society by John W. Atkinson © 1958 by Litton Educational Publishing, Inc. Reprinted with permission of Van Nostrand Reinhold Company).

Further, since all groups worked on the same set of problems, the intensity of their performance is directly reflected by the score they obtained on the task. Figure 3 presents the data obtained by Atkinson (1958) in the two incentive conditions used. Both incentive groups evidence the inverted- U function predicted in the previous section. On the other hand, the observed elevation of the performance curve due to increased incentive is not explicable in terms of the analysis presented here. This problem receives further attention in the General Discussion section below.

When the formal derivation of Atkinson's result is stated in ordinary language, it acquires a compelling simplicity. Performance level is relatively low when the task is perceived to be very easy because subjects generally believe that great effort is not necessary in order to succeed. When the task is thought to be very difficult, performance level is again low, this time because subjects believe that great effort will be of no avail. On the other hand, it is just when the task is perceived to be neither extremely easy nor extremely difficult that effort is seen as the crucial determinant of success or failure. Thus, performance level is highest at tasks of intermediate perceived difficulty.

More direct evidence for the analysis presented here comes from a recent experiment by Weiner, Heckhausen, Meyer, and Cook (1972). These authors asked subjects to indicate how much effort they would expend if they were to undertake any of a number of tasks varying in difficulty. The data show that the relationship between task difficulty and intended effort expenditure is an inverted-U function. This finding directly verifies the relationship between intended effort and perceived difficulty predicted by the attributional theory of performance.

INTERACTION OF PERCEIVED DIFFICULTY AND PERCEIVED ABILITY

Concept of Ability

It is likely that a person believes that less effort is necessary for success at a task if he thinks himself very able than if he believes that his ability for the task is relatively low (Heider, 1958, p. 111). Thus, the fourth Heiderian attributional variable of preceived ability is also a determinant of performance intensity. The attributional theory of performance is capable of making a number or predictions concerning the relative performance of groups which differ in their perceived ability for the task at hand. Unfortunately, the behavioral characteristics of perceived-ability groups (as opposed to real-ability groups) have received little explicit attention. The theoretical predictions made below are thus not directly verifiable in the existing literature.

However, a recent finding relates the concept of perceived ability to a frequently studied individual difference variable. An experiment by Kukla (1972, Experiment 1) has shown that subjects high in resultant achievement motivation, which is defined as the strength of achievement needs minus the strength of anxiety, report a higher estimate of their ability than low resultant achievers after having worked on an identical task. Yet the task outcome in this experiment was not determined by skill at all, so that this finding is not explicable by supposing that the high resultant achievers were in fact higher in ability. Rather, it seems that the higher resultant achieving subject is characterized by the general disposition to attribute relatively high ability to himself, while the low resultant achiever has a generalized tendency to attribute relatively little ability to himself. A related result of Atkinson, Bastian, Earl, and Litwin (1960), who found that subjects high in achievement needs exceeded low achievers in their estimates of how well they expected to do at a novel task, strongly suggests that this perceived-ability differential is due to a generalized tendency rather than being a task-specific response.

In view of these findings, the large number of published investigations dealing with resultant achievement motivation becomes relevant to the present theory's predictions concerning the effects of perceived ability. Specifically, any hypothesis concerning the behavioral differences expected between high and low perceived-ability groups immediately generates the second hypothesis that the same differences are to be expected between high and low resultant achievement groups. This second hypothesis is invariably more likely to have been tested by prior research.

There is, of course, an implicit assumption in the above argument: it is necessary to suppose that the perceived-ability differential is the only motivationally significant difference between the two resultant achievement groups. This view differs fundamentally from the more well-known theories of achievement motivation and of anxiety (Atkinson, 1964; Mandler & Sarason, 1952; Spence, 1958). The general position taken here, that individual differences in achievement-related behavior are due to differences in the subjects' cognitive appraisals of task situations, has been

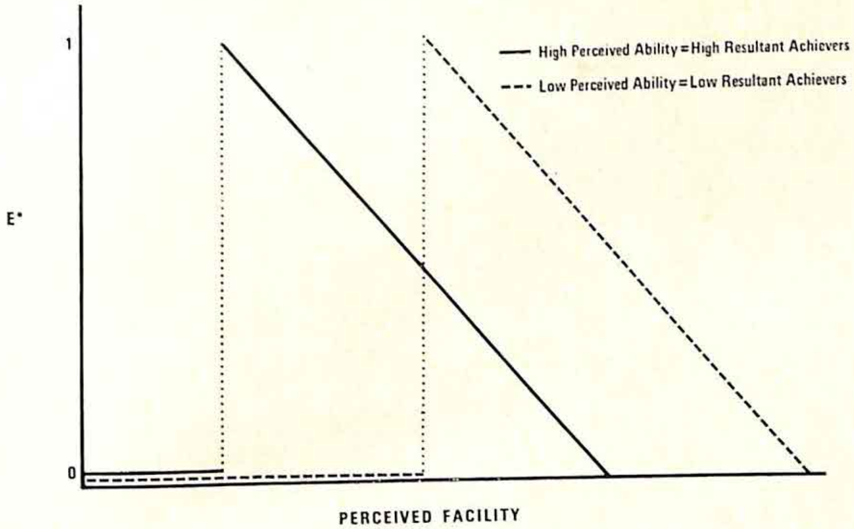


FIG. 4. Theoretical relationship between perceived task facility and E^* for individuals high and low in perceived ability for the task.

argued for at length by Weiner, Frieze, Kukla, Reed, Rest, and Rosenbaum (1972). Further, Kukla (1972, Experiments 2 and 3) has shown that such cognitive appraisals definitely influence both the direction and the intensity of behavior. Thus, equating resultant achievement motivation with a generalized disposition to perceive one's ability as high or low has some prima facie justification. This point of view is made still more attractive by the fact, demonstrated below, that it results in a theoretical analysis which adequately explains a considerable body of data concerning the relative performances of high and low resultant achievement groups.

Theoretical Analysis

The following postulate concerning perceived ability is adopted: For a task which two subjects perceive to be equally difficult, the subject who believes his ability to be higher will also expect that a smaller amount of effort is required for him to succeed.

Consider now a task X of arbitrary perceived difficulty. Let E_h^* be the intended effort level of a subject S_h who perceives his ability for the task to be high, and let E_l^* be the intended effort level of a subject S_l who perceives his ability for X to be low. Equivalently, S_h and S_l can be subjects high and low in resultant achievement motivation. Now E_h^* and E_l^* are the smallest effort levels which each subject expects will lead to success. Thus, by the assumption of the previous paragraph, $E_h^* < E_l^*$; that is, a subject who perceives his ability to be high will intend to expend less effort on the same task than a subject who believes his ability is low.

For any task perceived as easier than X, both subjects will decrease their intended effort level, S_l maintaining a higher E^* than S_h. Similarly, if a task is thought to be more difficult than X, both subjects will increase their E^*, and again E_l^* will remain higher than E_h^*. However, there is a maximum possible level for E^*. As tasks of increasing perceived difficulty are considered, E_l^* will reach this maximum sooner than will E_h^*, that is, S_l will more quickly conclude that the task requires his greatest possible exertion. For any task whose perceived difficulty is greater than this, S_l will believe that success is not possible with any level of exertion, and so by the principle of least effort, he will decide not to exert himself at all. On the other hand, S_h will not yet have reached his maximal E^*, so that for the same task at which S_l's intended effort drops to zero, his intended effort will continue to rise.

If perceived difficulty becomes still greater, there of course comes a point at which S_h too concludes that success is not possible and ceases to exert himself.

Figure 4 depicts one possible configuration of relationships among the variables of perceived ability (or resultant achievement motivation), perceived difficulty, and intended effort. A number of the features of the figure are arbitrary, as for example the parallel slopes of the two curves. The derived properties of the figure are (*a*) the existence of a region of high perceived facility where $E_h^* < E_l^*$; (*b*) the existence of a region of intermediate perceived facility where $E_h^* > E_l^*$; and (*c*) the existence of a region of very low perceived facility where $E_h^* = E_l^* = 0$.

Figure 4 depicts individual performance curves. Group curves should approximate two overlapping inverted-U functions. Again, since behavioral intensity is a positive linear function of intended effort, Figure 4 also depicts the functional relationship between perceived facility and intensity.

Verifying Data

Partial verification for the above predictions comes from a well-known series of experiments undertaken by Spence and his colleagues (Spence, Farber, & McFann, 1956; Spence, Taylor, & Ketchel, 1956; Taylor & Chapman, 1955). These authors found that subjects high in anxiety performed better at (objectively) easy tasks but worse at (objectively) difficult tasks than did low-anxiety subjects. If we assume that with continued work on the task and in the absence of contradicting information, perceived difficulty comes to approximate objective difficulty, these results are in accord with our derivation. Figure 4 shows that when perceived task facility is high, low resultant achievers (who are on the whole high in anxiety) select a greater E^* than do high resultant achievers (who tend to be low in anxiety). Conversely, when the task is perceived to be relatively more difficult, the intended effort level of high resultant achieving subjects surpasses that of low resultant achievers.

It can be shown that Spence's explanation of these results is incompatible with other data which are also in accord with the present theory. Spence's account rests on the identification of task difficulty with the relative strength of the correct "habit" in the subject's habit-family hierarchy. Thus, for Spence the observed interaction is interpreted as signifying a relationship between performance and the *objective* difficulty level of the task. On the other hand, the attributional theory of performance explains the same results as due to the subject's selection of an effort level based on his *perception* of the task's difficulty, regardless of whether this perception is veridical. An important experiment by Weiner (1966) demonstrates that the present theory is correct on this point. Weiner experimentally separated objective from perceived task difficulty by leading subjects to believe that they were succeeding at what was an objectively difficult task and that they were failing at what was an objectively easy task. Under the assumption that continued success leads to the perception that the task is easy whereas continued failure leads to the perception that the task is difficult, the two groups were so constituted that one worked on an objectively easy but subjectively difficult task and the other on an objectively difficult but subjectively easy task. Weiner found that subjects high in anxiety were superior to low-anxiety subjects in the second of these conditions, but were inferior in the first. Thus Spence's data concerning the effects of task difficulty on performance were reversed.[4] Yet, according to Spence's view, since the habit-family hierarchy was not affected by these experimental operations, the results should also not have been altered. To be sure, Spence recognized that the experience of success or failure influences performance, but his theory postulates that this influence is mediated by a change in anxiety level. Such a change could either improve or diminish the performance of all subjects, depending on the difficulty of the task; but it could not result in the reversal reported by Weiner of the usually obtained interaction between difficulty and anxiety.

---

[4] Weiner and Schneider (1971) have reported a full replication of this finding.

More complete confirmation of the present analysis comes from the experiment by Atkinson discussed in an earlier section (Atkinson, 1958, Ch. 10). In that experiment, level of performance was measured as a function of perceived difficulty, the latter quantity being conceived as varying with the proportion of subjects who could win a prize. Figure 5 shows the task performance data plotted separately for high and low achieving groups in the $1.25 incentive condition. Although the difference in performance between these two groups is not statistically significant for any single level of perceived difficulty, one is struck by the overall similarity between these empirical curves and the theoretically derived curves of Figure 4.

INTERACTIONS BETWEEN PERCEIVED ABILITY, INITIAL PERCEIVED DIFFICULTY, AND SUCCESS AND FAILURE

A number of specific predictions can be derived for situations in which subjects are led to have an initial estimate of the task's difficulty and subsequently also experience a series of successes or failures. It is in such situations especially that the present analysis is capable of generating untested hypotheses. Only the more important of these hypotheses are outlined below. Their derivation follows directly from a previously used assumption concerning the effects of success and failure on perceived difficulty: A series of success experiences (whether veridical or not) causes perceived difficulty to decrease, while a series of failure experiences (whether veridical or not) causes perceived difficulty to increase.[5] The analysis below refers to high and low resultant achievement motivation with the understanding that these terms are interchangeable with high and low perceived ability, respectively.

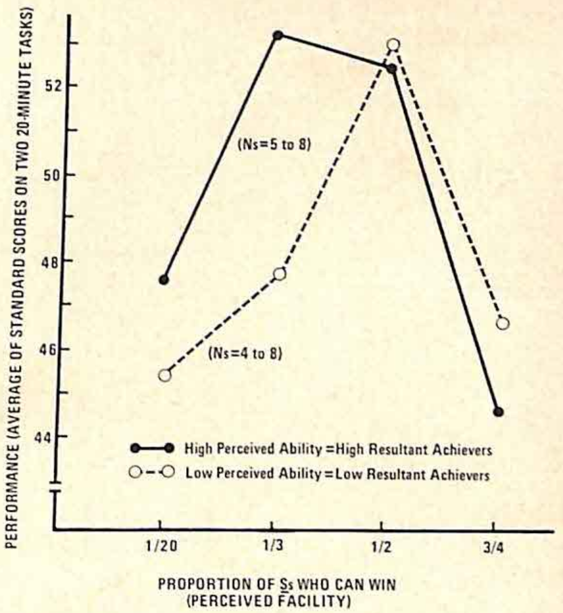


FIG. 5. Observed relationship between intensity of performance and perceived task facility for high and low achievement groups (adapted from Atkinson, 1958, Ch. 10, p. 295. From Motives in Fantasy, Action and Society by John W. Atkinson © 1958 by Litton Educational Publishing, Inc. Reprinted with permission of Van Nostrand Reinhold Company).

We first define a task as (subjectively) *easy*, if it lies within the range of perceived difficulties for which both the high resultant achiever's and the low resultant achievers' effort functions increase with increased difficulty (range X to Y in Figure 6). A task of *intermediate* (subjective) difficulty is one which lies within the range Y to Z in the figure; by definition then, an intermediate task is one at which a high resultant achiever believes he is capable of success, but which a low resultant achiever considers beyond his capacities. Finally, a task will be said to be (subjectively) *difficult* if it is perceived to be more difficult than Z, that is, if both high and low resultant achievers expect to fail at it with even maximal effort.

Suppose that all experimental groups work on tasks of equal objective difficulty and that they all have equal average ability for the tasks; thus, performance level can be expected directly to reflect effort expenditure. With the definitions and assumptions introduced above, the following hypotheses follow immediately.

---

[5] Alternatively, success may lead to an increase and failure to a decrease in perceived ability. Although there are testable differences between these two possibilities, the hypotheses discussed here follow equally from the assumption of a change in perceived difficulty or a change in perceived ability.

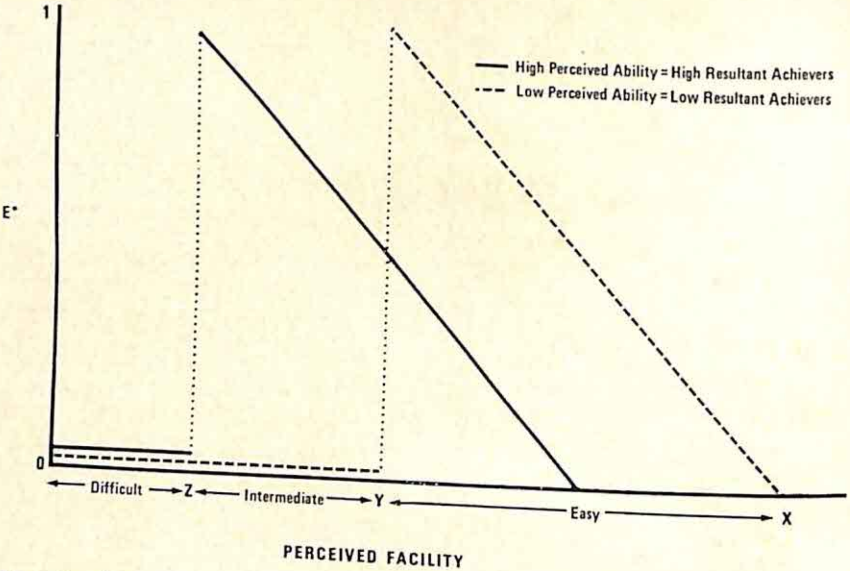


FIG. 6. Definitions of subjectively easy, intermediate, and difficult tasks in terms of the relative effort functions of high and low resultant achievers.

1. Given a task initially perceived as easy, the experience of continuous failure (whether veridical or not) will result in an increase in the performance level of all subjects. If failure continues to be forthcoming, all subjects will reach a performance peak after which their performance will deteriorate. The performance of low resultant achievers will begin to deteriorate sooner than that of high resultant achievers.

2. Given a task initially perceived as difficult, the experience of continuous success (whether veridical or not) will result in an increased performance level for all subjects. If success continues to be forthcoming, all subjects will eventually reach a performance peak and then exhibit a decrease in performance level. The performance of high resultant achievers will begin to decrease sooner than that of low resultant achievers.

3. Given a task initially perceived as intermediate in difficulty, a series of successes (whether veridical or not) will result in an improved performance on the part of low resultant achievers but in a diminished performance by high resultant achievers. Conversely, a series of failures (whether veridical or not) will lead to an increase in the performance of high resultant achievers but a decrease in the performance of low resultant achievers.

There are data in the literature which are compatible with the above derivations. But since experiments have not generally assessed or systematically varied initial perceived difficulties when success and failure manipulations have been used, it is not possible to claim that they provide strong verification for the theory. Thus, several studies have found that a series of failures (with an unspecified initial perceived difficulty) resulted in an improved performance among subjects low in anxiety (and so high in resultant achievement motivation) but in a diminished performance among highly anxious (low resultant achieving) subjects (Katchmar, Ross, & Andrews, 1958; Lucas, 1952; Zeller, 1951). These findings are in accord with the present theory under the assumption that the tasks used were initially of intermediate perceived difficulty. Similarly, there are studies reporting that a series of successes leads to increments in performance among high-anxiety subjects but to performance decrements among subjects low in anxiety (Child & Whiting, 1950; Mandler & Sarason, 1952). These findings too are in accord with the attributional theory of performance if the tasks used were initially of intermediate perceived difficulty. Nevertheless, there is evidently a need for empirical investigations of the combined effects on performance of initial perceived difficulty and success–failure experiences.

GENERAL DISCUSSION

Comparisons with Atkinson's Theory of Achievement Motivation

The theoretical analysis most closely related to the present approach, in terms of the body of data it attempts to explain, is the theory of achievement motivation formulated and several times revised by Atkinson and his colleagues (Atkinson, 1957; Atkinson & Cartwright, 1964; Atkinson & Feather, 1966; Weiner, 1970). Atkinson's theory also adopts the conceptual foundation of expectancy theory, but adds to this a number of distinctive assumptions concerning the interdependence of utility and P_s in achievement-oriented situations, the effects of the personality dispositions of need for achievement and anxiety, and the relationship between *SEU* (called "motivational tendency") and performance level. These assumptions need not be reviewed here.

Several of the resulting derivations are similar to those of the attributional theory of performance. Thus, Atkinson's theory also predicts, at least for subjects high in resultant achievement motivation, that the relationship between performance level and perceived difficulty (which Atkinson equates with P_s) will be an inverted-U function. On the other hand, achievement theory suggests that the same function for low resultant achieving subjects will be a non-inverted-U curve, whereas the present theory predicts another inverted-U curve translated to the right (see Figure 4). While the data discussed in this paper (Atkinson, 1958, Ch. 10) support the present analysis, this support is not as conclusive as one might wish. First, the experiment cited employed female subjects, and there is some question about whether the measure of achievement motivation used is valid for female populations (Veroff, Wilcox, & Atkinson, 1953). Second, Karabenick and Youssef (1968) have reported an instance of the non-inverted-U performance curve for low achievers expected by achievement theory. Thus, the issue of whether achievement theory or the present attributional theory is correct on this point rests on whether the Atkinson (1958) result or the Karabenick and Youssef (1968) result proves to be reliable.

A point of divergence between the two theories which at present favors achievement theory has previously been mentioned. The increase in performance level when the utility (incentive value) of success is increased, as observed in Atkinson (1958), is not explicable by the present theory. On the other hand, Atkinson's theory which postulates (with some qualifications, e.g., Atkinson & Feather, 1966, pp. 318–324; Atkinson & Reitman, 1956) that performance level increases as *SEU* increases predicts precisely this result. If this finding should prove reliable, it indicates a need for future revisions of the present theory. Yet, the question of whether performance is generally an increasing function of incentive is far from conclusively settled. Even among experiments using animal subjects, it is not always the case that "response strength" (performance level) is observed to be an increasing function of "amount of reinforcement" (utility, incentive). Although the classical findings of Crespi (1942) and Zeaman (1949) indicate such a relationship, Snyder and Hulse (1961) reported an experiment in which neither the total volume of water available to thirsty rats after each trial nor the number of licks they were allowed to make affected the speed with which they ran down the alley after the training trials had been completed. Also, Schrier (1958) has shown that monkeys given one, two, four, or eight food pellets for a correct response at a discrimination task do not differ in the total percentage of correct responses made; however, in an additional group which experienced all four levels of reward in the pretraining period, performance proved to be a function of the amount of reward used during training. Thus, it is possible that the relationship between utility and performance is rather more complex than it is usually conceived. At present, neither

Atkinson's theory nor the attributional theory of performance are in a position to explain the precise manner in which performance level varies with utility.

The fact, suggested by Atkinson's (1958) experiment, and strongly confirmed by the experiments of the Spence group, that low resultant achievers perform absolutely better than high resultant achievers when the task is perceived to be easy favors the present theory. The attributional theory of performance directly predicts this result, whereas Atkinson's theory, as well as its later revisions, would seem to predict that a group of subjects low in resultant achievement motivation could never exceed a high resultant achieving group in performance. Yet this point too cannot be said to be conclusive. For if achievement theory adds the assumptions (*a*) that subjects labeled low resultant achievers in most experiments are still relatively high in resultant achievement motivation, and (*b*) that the higher resultant achievement group perceives the task as being easier than does the lower resultant achievement group, then it follows that the lower resultant achievers will perform better than the higher resultant achievers. These two assumptions, although post hoc, are not unreasonable. Since the subjects used in experiments are generally college students, it may well be that those persons extremely low in resultant achievement motivation are excluded from consideration. Further, there is some experimental evidence that high and low resultant achievers diverge in their perception of task difficulty in the way required for this explanation (McClelland, 1961; Pottharst, 1955). However, Atkinson does not consistently invoke this latter assumption in experimental situations where it would substantially modify his theoretical predictions (see, for example, Atkinson & Litwin, 1960).

Finally, the findings that high resultant achievers improve after failure but do worse after success, whereas low resultant achievers do worse after failure but improve after success, can be equally well explained both by the present theory and by achievement theory.

In sum, it is clear that a definitive comparison of the adequacy of the two theories must await further research.

Theoretical Extensions

The arguments in this paper have been undertaken with severely limiting assumptions. In particular, the assumptions that each effort level is associated with certainty to a specific outcome, and that each task results in one of only two possible outcomes, restrict the range of applicability of the theory to a far greater degree than is necessary. Derivations for skill situations in which chance also plays a role, or where tasks have an indefinitely large number of possible outcomes, pose no serious difficulties. The basic principles of the attributional theory of performance were initially formulated for these more general cases, and it remains only to draw out the implications of these principles.

The familiar performance theories of Spence, of Mandler and Sarason, and of Atkinson are also restricted to dichotomous-outcome tasks, the two outcomes being conceived as success and failure by Atkinson, and as the emission of either a correct or an incorrect response by the other two theories. It does not seem as if this limitation is any more inherent in the conceptual apparatus of these theories than it is in the present theory's. Indeed, Logan's (1956) micromolar analysis consists in part of an extension of the Hull-Spence approach to multiple-outcome situations. However, it is more doubtful whether the classical performance theories can, without significant revisions, conceptualize task situations where both skill and chance determine outcome.[6] They thus share the shortcoming of having insufficient means for the description of a wide variety of perceived task structures which differentially affect performance.

On the other hand, Atkinson's theory contains an explicit theory of choice and thus is relevant to a large body of data to which the present theory does not address itself. It is however evident that the attributional theory of performance provides a conceptual foundation in terms of which a theory of choice can readily be formulated. For example, one can adopt a modified version of the choice principle employed by Atkinson, that the alternative with the greatest SEU is likeliest to be selected. In the attributional theory of performance, SEU is a function of effort. One may then stipulate that, of a set of available alternatives $X_1, X_2, \cdots X_n$, subjects first calculate for each X_i the maximal SEU possible for any possible level of E, and then select that action whose maximal SEU is greater than any other action's. An elaboration of such a choice principle within the attributional theory of performance will provide new grounds for testing the relative merits of the present approach and achievement theory.

---

[6] Feather (1967) has discussed the sort of theoretical revision necessary for Atkinson's theory to be capable of dealing with mixed skill–chance situations.

REFERENCES

ADAMS E. W., & FAGOT, R. A model of riskless choice. *Behavioral Science*, 1959, **4**, 1–10.

ATKINSON, J. W. Motivational determinants of risk-taking behavior. *Psychological Review*, 1957 **64**, 359–372.

ATKINSON, J. W. Towards experimental analysis of human motivation in terms of motives, expectancies, and incentives. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society*. Princeton, N. J.: Van Nostrand, 1958.

ATKINSON, J. W. *An introduction to motivation*. Princeton, N. J.: Van Nostrand, 1964.

ATKINSON, J. W., BASTIAN, J. R., EARL, R. W., & LITWIN, G. H. The achievement motive, goal setting, and probability preferences. *Journal of Abnormal and Social Psychology*, 1960, **60**, 27–36.

ATKINSON, J. W., & CARTWRIGHT, D. Some neglected variables in contemporary conceptions of decision and performance. *Psychological Reports*, 1964, **14**, 575–590.

ATKINSON, J. W., & FEATHER, N. T. (Eds.), *A theory of achievement motivation*. New York: Wiley, 1966.

ATKINSON, J. W., & LITWIN, G. H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *Journal of Abnormal and Social Psychology*, 1960, **60**, 52–63.

ATKINSON, J. W., & REITMAN, W. R. Performance as a function of motive strength and expectancy of goal attainment. *Journal of Abnormal and Social Psychology*, 1956, **53**, 361–366.

CHILD, I. L., & WHITING, J. W. M. Effects of goal attainment: Relaxation versus renewed striving. *Journal of Abnormal and Social Psychology*, 1950, **45**, 667–681.

CRESPI, L. P. Quantitative variation of incentive and performance in the white rat. *American Journal of Psychology*, 1942, **55**, 467–517.

EDWARDS, W. The theory of decision making. *Psychological Bulletin*, 1954, **51**, 380–417.

EDWARDS, W. Behavioral decision theory. *Annual Review of Psychology*, 1961, **12**, 473–498.

EDWARDS, W., & TVERSKY, A. (Eds.) *Decision making*. Baltimore: Penguin, 1967.

FEATHER, N. T. Subjective probability and decision under uncertainty. *Psychological Review*, 1959, **66**, 150–164.

FEATHER, N. T. Valence of outcome and expectation of success in relation to task difficulty and perceived locus of control. *Journal of Personality and Social Psychology*, 1967, **7**, 372–386.

FRIEZE, I., & WEINER, B. Cue utilization and attributional judgments for success and failure. *Journal of Personality*, 1971, **39**, 591–605.

HEIDER, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.

KARABENICK, S. A., & YOUSSEF, Z. I. Performance as a function of achievement motive level and perceived difficulty. *Journal of Personality and Social Psychology*, 1968, **10**, 414–419.

KATCHMAR, L. T., ROSS, S., & ANDREWS, T. G. Effects of stress and anxiety on performance of a complex verbal-coding task. *Journal of Experimental Psychology*, 1958, **55**, 559–563.

KELLEY, H. H. Attribution theory in social psychology. *Nebraska Symposium on Motivation*, 1967, **15**, 192–240.

KUKLA, A. Attributional determinants of achievement-related behavior. *Journal of Personality and Social Psychology*, 1972, **21**, 166–174.

LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.

LEWIN, K. *The conceptual representation and the measurement of psychological forces*. Durham, N. C.: Duke University Press, 1938.

LOGAN, F. A. A micromolar approach to behavior theory. *Psychological Review*, 1956, **63**, 63–73.

LUCAS, J. D. The interactive effects of anxiety, failure and interserial duplication. *American Journal of Psychology*, 1952, **55**, 59–66.

MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 1952, **47**, 166–173.

MCCLELLAND, D. C. *The achieving society*. Princeton, N. J.: Van Nostrand, 1961.

MOSTELLER, F., & NOGEE, P. An experimental measure of utility. *Journal of Political Economy*, 1951, **59**, 371–404.

POTTHARST, B. C. The achievement motive and level of aspiration after experimentally induced success and failure. Unpublished doctoral dissertation, University of Michigan, 1955.

ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.

ROTTER, J. B. The role of the psychological situation in determining the direction of human behavior. *Nebraska Symposium on Motivation*, 1955, **3**, 245–269.

SCHRIER, A. M. Comparison of two methods of investigating the effects of amount of reward on performance. *Journal of Comparative and Physiological Psychology*, 1958, **51**, 725–731.

SNYDER, H. L., & HULSE, S. H. Effect of volume of reinforcement and number of consummatory responses on licking and running behavior. *Journal of Experimental Psychology*, 1961, **61**, 474–479.

SPENCE, K. W. A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations. *American Psychologist*, 1958, **13**, 131–141.

SPENCE, K. W., FARBER, I.E., & MCFANN, H. H. The relation of anxiety (drive) level to performance in competitional and noncompetitional paired-associates learning. *Journal of Experimental Psychology*, 1956, **52**, 296–305.

SPENCE, K. W., TAYLOR, J., & KETCHEL, R. Anxiety (drive) level and degree of competition in paired-associates learning. *Journal of Experimental Psychology*, 1956, **52**, 306–310.

TAYLOR, J., & CHAPMAN, J. P. Paired-associate learning as related to anxiety. *American Journal of Psychology*, 1955, **68**, 671.

TOLMAN, E. C. *Purposive behavior in animals and men*. New York: Century, 1932.

TOLMAN, E. C. Principles of purposive behavior. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959.

VEROFF, J., WILCOX, S., & ATKINSON, J. W. The achievement motive in high school and college age women. *Journal of Abnormal and Social Psychology*, 1953, **48**, 108–119.

WEINER, B. Role of success and failure in the learning of easy and complex tasks. *Journal of Personality and Social Psychology*, 1966, **3**, 339–344.

WEINER B. New conceptions in the study of achievement motivation. In B. Maher (Ed.), *Progress in experimental personality research*. Vol. 5. New York: Academic Press, 1970.

WEINER, B., FRIEZE, I., KUKLA, A., REED, L., REST, S. A., & ROSENBAUM, R. M. Perceiving the causes of success and failure. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior*. New York: General Learning Press, 1972.

WEINER, B., HECKHAUSEN, H., MEYER, W., & COOK, R. Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. *Journal of Personality and Social Psychology*, 1972, **21**, 239–248.

WEINER, B., & KUKLA, A. An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 1970, **15**, 1–20.

WEINER, B., & SCHNEIDER, K. Drive versus cognitive theory: A reply to Boor and Harmon. *Journal of Personality and Social Psychology*, 1971, **18**, 258–262.

ZEAMAN, D. Response latency as a function of the amount of reinforcement. *Journal of Experimental Psychology*, 1949, **39**, 466–483.

ZELLER, A. F. An experimental analogue of repression: III. The effects of induced failure and success on memory measured by recall. *Journal of Experimental Psychology*, 1951, **42**, 32–38.

(Received December 21, 1971; revision received May 15, 1972)

Psychological Review
1972, Vol. 79, No. 6, 471–486

THE THEORY OF COGNITIVE RESIDUES: A NEW VIEW OF FANTASY

DAVID A. ROSENBAUM[1]

Swarthmore College

Cognitive *residues* are structures that were absorbed or displaced in development. Because residues remain "intact," they seek expression or "reification." Residual reification contradicts reification of predominant structures and so it requires a special innocuous context which is provided in fantasy. Residual fantasy is self-regenerative within and among individuals. Naive perceivers assimilate residual fantasy expressions of their elders, thus forming constructs which later become residues. Residual fantasy is transmitted over generations and is abetted by the growth of technology. Use of hallucinogenic drugs may occur when technological means of residual expression become inaccessible. To reduce formation of residues, parents may indicate the special context of fantasy and prescribe a wide variety of clearly fantastic exposures for their children. The importance of the theory is considered.

The origins of fantasy have been explained in terms of three models. The first has been formalized and widely promulgated. The second has not been formalized but has apparently been intuitively accepted by the social scientific community. The third has not yet been placed in the public domain. It will be the aim of this paper to formalize the second model and to introduce and elaborate the third.

Freud's theory of *wish fulfillment* represents the first model. According to Freud, fantasy represents the disguised expression of repressed sexual drives. The purpose of the expression is gratification of the drives. Because the gratification is only temporary, a renewed cathexis becomes necessary when gratification wears off. For our purposes, the main consequences of this "hydraulic" model of fantasy have been the symbolic interpretation of dreams, projective testing, and the psychoanalytic interpretation of visual art, as it has been discussed by Freud (1914, 1916), Devereux and LaBarre (1961), Ehrenzweig (1962), and Waelder (1965).

The second approach to fantasy has been embodied in what can be called an *exploratory* model. In the view of this model, fantasy represents the "reification" of newly evolved, though still tentative ideas. Reification will be explained in a subsequent section of this paper. For present purposes, therefore, the exploratory model will be elucidated by way of example. The clearest example is science fiction. The author of science fiction invents unfamiliar situations which demand novel responses. He "tests" his new ideas for the possibility of their actualization in the environment by exploring their consistency with each other and with the physical world. Once a possibility of occurrence is established, the implications of this possibility are explored. A second example of exploratory fantasy is the vicarious assumption of roles. As Sanford (1955), Parsons (1961), and Burton and Whiting (1961) have pointed out, children rehearse in fantasy the behaviors appropriate to roles they may wish to assume. Their covert assumption of these roles lets them test wide varieties of behavior without running the risk of being punished for violating the rules of conduct prescribed for children. It should be obvious that the interpretations of both of these examples presuppose a great deal about cognitive mechanisms and motivation. The speculation attached to the examples will hopefully be substantiated in a satisfactory way as the discussion progresses.

The third model posits fantasy as an expression of cognitive structures modified in development. Fantasy is seen here as an expression (or reification) of cognitive structures which remain "intact" in the mind after being displaced or absorbed by other structures. The modified, remaining structures, or what I have called cognitive *residues,* will occupy the main focus of our attention in this paper. Considerable attention will be devoted to two major implications of the residual model: (*a*) Propagation of fantasy in culture causes a long-term spiralling of fantasy expression which is abetted by the growth of technology and which has recently been manifested dramatically in the widespread use of drugs. (*b*) At critical periods of development, it may be possible to prevent the formation of large numbers of cognitive residues by indicating differences between "reality" and fantasy in order to help reduce the "need" for fantasy in the individual's life. Both of these implications are critically examined to determine their logical and empirical validity.

---

[1] Requests for reprints should be sent to David A. Rosenbaum, Swarthmore College, Swarthmore, Pennsylvania 19081.

In completing this outline of the three models of fantasy, it should be stated that the models are not thought to be mutually exclusive. No logical inconsistency rests in the statement of their coexistence or in their joint operation in man. In juxtaposing the three models, it has been implied that social scientists should be able to see in fantasy the expression of epistemic curiosity and conservatism as well as sexuality. Where this broadened viewpoint may now be hindered is in an understanding of what is meant by "epistemic curiosity and conservatism" and "reification." The next section should help to clarify these concepts.

REIFICATION

Reification can be defined here as the tendency to seek the fulfillment of expectations in as wide a range of experience as expectations allow. The range of experience can be said to encompass the whole of the interactional field between man and environment, embodying what Klein (1958) has called the "meaning-inducing field" and what Hallowell (1955) has called the "behavioral environment."

Levels of Activity

Reification operates on an active and less active (though not passive) level. The cognitive processes of each of the levels can be related to Kelly's (1963) functions of "prediction" and "control" and to Piaget's (e.g., 1952b, 1957, 1970) functions of "accommodation" and "assimilation."

The less active level is characterized by expectation alone. The individual predicts events without affecting the environment in a way that will maximize the chances of fulfilling his predictions. If predictions come true, correctly anticipated events are assimilated into the cognitive structure. If predictions do not come true, some accommodation becomes necessary. The existence of this level of reification is perhaps more didactically appealing than theoretically or practically possible. By virtue of the psychological pervasiveness of the meaning-inducing field, any prediction necessarily affects this field in an active way. Furthermore, receptivity to environmental phenomena is shaped by expectations, as research in perception has demonstrated (e.g., Wallach, 1956; Zangwill, 1937).

The more active level of reification is probably more realistic. Here the individual controls his environment by affecting events (and objects) in a way that will maximize the chances of fulfilling expectations about them. The individual not only assimilates events which happen to conform to his predictions but actively shapes these events (as much as possible) to guarantee their conformity. Perhaps the most obvious example of this process is goal-directed activity. As outlined by Miller, Galanter, and Pribram (1960) in their TOTE sequence for plans, the organism continues to act until the goal of its action has been achieved. Spiro (1964) has indicated that the presence of goals implies cognitive assumptions concerning form, meaning, and the *efficacy* of action. Klein (1958) has shown that the terminus of goal-directed action must be seen as "'adequate' to the initiating intention [p. 95]." In light of the present discussion, "adequacy" can be defined as the accomplishment of reification. An action is seen as adequate when expectations about its results are realized. In the goal-directed activity of manufacturing artifacts, for example, this activity is judged adequate when the artifacts have been completed in their expected form.

Goal-directedness is not the only instance of the more active type of reification. Social gregariousness, as Schachter (1959) has suggested, serves a function of self-evaluation. Motor activity causes neural excitation which provides sensory feedback for the organism—Held's (1965) "reafference" and Troland's (1932) more general "retroflex action." Although it is possible that self-evaluation and sensory feedback may, in fact, exemplify goal-directedness,[2] both of these informational sources serve to confirm or contradict anticipations about social and motor effectiveness. Future behavior stands a better chance of being confirmed as operatively (or functionally) successful once evaluative and feedback information is used to modify behaviors which contravene expectations.

Epistemic Confidence

The definition of reification includes a proviso for a kind of "epistemic confidence": Reification operates in "as wide a range of experience as expectations allow"; stated somewhat differently, the breadth of experience made amenable to prediction is itself predicted and thereby limited. The extent of limitation is determined by predictive "necessity" and "risk." Predictive necessity denotes the need to predict events whose occurrence determines the fulfillment of basic maintenance requirements (e.g., biological subsistence) of the reifying apparatus. Risk denotes the expansibility of the reifying ken. It is this latter ingredient which accounts for varying ranges of reification. Since the ability or inability to predict events must itself be predicted, the individual regulates his predictive activity according to his "confidence" in his predictive capacities. His confidence is largely shaped by his previous experience, especially by the informational content of past evaluations and feedback. Depending on the extent of his confidence, he may exhibit more curiosity on one hand or more conservatism and avoidance on the other. Curiosity implies an ever-widening field of reification, while conservatism implies a stable and restricted reifying field. It is unlikely that either one of these tendencies ever exists without the other in the individual's cognitive framework, as Hunt (1961, p. 269) and Prentice (1964) have said. The coexistence of curiosity and conservatism is vital to the individual's successful adaptation in the environment. The two tendencies, respectively, ensure adjustment to novel situations and retention of capacities and concepts acquired in past experience.

Conflict

Reification is most clearly manifested after its own failure. When surprise follows certain events, it can be inferred that these events were not expected but that others were. It can be surmised that a conflict existed between what was expected and what actually occurred. This assumption has been supported by Charlesworth (1969) and has been applied empirically by Bower (1971). For our purposes, there is a twofold importance to the assumption. First, reactions to conflict can be used to test the presence and scope of the reification process. Second, the range of experience in which conflict is perceived demonstrates the pervasiveness of the reifying field.

Reaction to conflict extends from the physiological to the conceptual level. The physiological responses to conflict have been described extensively (Jacobson, 1926; Landis & Hunt, 1939; Skaggs, 1926; Strauss, 1929). Berlyne (1962, 1964) has indicated the "attention value" of these responses, much as Piaget (1952, 1957) has shown their "curiosity value" and Pavlov (Charlesworth, 1964) has proclaimed their values as "orienting" reflexes. The values ascribed to these responses have been determined from the perspective of their general function—reduction and, if possible, resolution of the conflicts which induced them. The specification of this function has been supported by empirical data, most notably that of Festinger (1957, 1964), and has been invoked as an essential input in the Piagetian model of development. It is not necessary to review the means of conflict reduction here, although Berlyne's (1964) article "Toward A Theory of Epistemic Behavior: Conceptual Conflict and Epistemic Curiosity" is highly commendable in this regard. It need only be said that the tendency to reduce or resolve conflict reflects the tendency to minimize the chances of the recurrence of unanticipated events. By reducing or resolving conflict, the individual reshapes his expectations about events which were formerly conflictive. In so doing, he allows for the expectation (or expectability) of these events and thereby ensures their reification in future experience.

---

[2] Ascribing goal-directedness to the two cases would be taking a considerable teleological risk. Nevertheless, the suggestions of Piaget and G. H. Mead, to name a few, raise at least interesting philosophical questions about such a possibility. Piaget (1954) has spoken of the "Construction of Reality" and Mead (1934) has insisted that the organism "determines the environment . . . the organism, then, is in a sense responsible for its environment [pp. 129–130]." Assuming this rather grandiose "imposition of arbitrary form," to use Holloway's (1969) phrase, any actions initiated in the environment could easily be seen as "subordinations of the world 'schema'" (drawing freely from Piaget). If "subordination" is an *intended* rather than a *retroactive* operation, all voluntary actions would, in fact, be goal-directed.

The range of experience in which conflict is perceived indicates the pervasiveness of the reifying field. Charlesworth (1969) has observed that conflicts are perceived between external stimuli, between internal and external stimuli, and between internal stimuli. The range of experience covered in Festinger's (1957) "cognitive dissonance," Abelson's (1958) "cognitive imbalance," and Berlyne's (1964) "conceptual conflict" seems to substantiate Charlesworth's claim. The inclusiveness of the conflictive classes outlined by the above authors corresponds to the inclusiveness of the reifying field as it has been conceived here.

Structural Basis of Expectation

In discussing reification, it becomes necessary to specify the nature of expectation. What is it that is expected? Is there an isomorphism between what is expected and what the content of an expectation reflects in the mind? An answer to these questions lies in structuralism: An expectation embodies the anticipated reflection of cognitive structures in the environment. Reflection denotes enactive, ikonic, or symbolic (Bruner, Olver, & Greenfield, 1966) extension. Cognitive structures have been defined by Piaget (1970) as "systematic wholes of self-regulating transformations [p. 44]." Lane (1970) has defined them as "sets of elements between which or between certain subsets of which relations are defined [p. 24]." Oldfield (1964) has defined structures in terms of circuital storage devices used in computers. Not to take issue with any of the above authors, for all of their definitions are compatible, cognitive structures—or what will also be called schemata and constructs—can also be thought of as explicit or implicit beliefs about the way things operate. Gravity, causality, and transitivity are representative beliefs. By identifying these beliefs, reification can be discussed in fairly definite terms. Assuming a belief in gravity, for example, (most) objects will be expected to remain earthbound. Should it occur that all familiar objects float off into space, the expectation about gravity would be contradicted and would have to be reshaped to account for this unprecedented event.

This simple-minded example illustrates "equilibration of assimilation and accommodation," which has served as the basic mechanism in Piaget's model of development. According to Piaget, cognitive modification is caused by the disequilibrating experience of conflict between structural anticipations and feedback. Some debate has arisen about the sensitivity to conflict among structures which coexist in a given stage of development. Structures at the same level, it has been argued, should not experience conflict if the (potential) anticipations of *later* stages are violated. Thus, acquisition of any stage of development should be final as long as anticipations derived from structures of the attained level are not contradicted. Flavell and Wohlwill (1969) have answered this argument. Drawing on data from empirical tests of the intertask consistency hypothesis implied in Piaget's notion of stage, they have recognized a need to identify a transition period for each stage of development. The transition period includes a period of "stabilization" and "consolidation" of new constructs. Flavell (1971) has also expanded the concept of transition to include "post-stage" development of stage-specific items. Post-stage development includes the gradual attainment of "functional maturity" in these items, that is, the highest level of their "evocability" and "utilizability." As a result of the transition period and post-stage development, developmentally heterogeneous structures exist together. This results in heterogeneous reactions to conflictive situations, which further results in conflicts among coexisting structures. To appease this conflict, further modification becomes necessary and development continues.

Conflict, then, provides strong evidence that structures rely on reification. Klein (1958) has gone further to say:

> Structures, once formed, seem to require a guaranteed supply of their appropriate level of stimulation—a particular frequency of contact with particular realms of objects and properties of objects. . . . Drastic alterations in the level of environmental stimulation—too low or too high—can cause adaptive structures to lose their effectiveness [p. 113].

Klein's conclusions are largely drawn from Hebb (1955), Piaget (1952), and Rapaport (1958). The "particular realms of objects and properties of objects" that he mentions can best be understood as reflections of the contents of cognitive structures in the environment. His assertion, if it is correct, adds an additional source of proof to the notion that structures rely on reification.

As a final note in this section on reification, the use of the word can itself be diversified. In view of the active nature of reification, to reify a cognitive structure is to predict and influence its reflection in the environment; a structure can be considered reified once its expected reflection in the environment has occurred; a structure that can be reified is reifiable and has reifiability.

RESIDUES

Neisser (1962) has described three types of structural accommodation. The first is absorption: old schemata are ingested into new, inclusive schemata and continue to have an effect on cognition, although they cannot be individually recalled. The second is displacement: "Part of the cognitive apparatus does not evolve, but continues to exist side by side with a new schema which assimilates the same environmental events in a different way [Neisser, p. 64]." The third is integration: "a step to another level of abstraction or understanding, in which outputs of the older models of processing are only part of a more comprehensive whole [Neisser, pp. 64–65]." Of the three accommodative methods, absorption and (particularly) displacement hold the most interest for us. Both of these methods yield modified structures which remain "intact" in the cognitive framework. The modified, remaining structures will henceforth be called cognitive residues.

Turning to the previous discussion on reification, the presence of residual structures implies the reifiability of these structures. Because of the reliance of structures on reification, the probability of residual reification is high. But residual reification seems doomed to fail. Residues become residues because they are not reified. They are displaced or absorbed by new structures which *are* reified and which, consequently, can become predominant in the cognitive framework. As long as these new constructs remain predominant, it can be assumed that they find continual reification in the environment and that residues do not. This means that if residues are to be reified, they must be reified in some environmentally anomalous way. Furthermore, their reification should not be reprehensible to predominant structures but should at least be granted a special context to prevent contradiction of the new structures. In this connection, the author would like to propose that such a special context may be provided in fantasy. Through fantasy, cognitive residues may accomplish reification.

The following examples will help to illustrate what is meant. All of the situations have actually been observed.

[Example 1]. A magician passes an object behind a handkerchief, then snatches the handkerchief away to reveal that the object has disappeared. The act is performed before an audience of adults and adolescents.

Piaget (1954) has shown that children in the sensorimotor period of development (ages 0–1½ years) only gradually come to understand the permanence of hidden objects. Thus it is possible to interpret the magician's display as a means of reifying residues which were formed in the minds of members of the audience during development of the object concept.

[Example 2]. A character in an animated cartoon runs over the edge of a cliff, stops in midair, stands still for a few seconds, looks down, realizes that he should fall, and finally does fall. The cartoon is watched by children and adults in a movie theater.

Two interpretations can be made, the first in reference to the sensorimotor period, the second in reference to children's egocentric beliefs. It can be claimed, first of all, that the cartoon reifies sensorimotor structures which (incidentally) allowed for the suspension of objects in midair. The sensorimotor child who has not yet grasped the object concept is:

> limited to pursuing, more or less correctly, with his eyes or with his hand the trajectory delineated by the movements of accommodation peculiar to the immediately preceding perception . . . [Piaget, 1954, p. 19].

Thus when Piaget drops an object in front of such a child, the child's visual search for the object is centered on Piaget's hand and on the space around Piaget's hand. Assuming that the cognitive structure which allowed for this kind of visual search remains, the animated cartoon can be viewed as a reification of this residue. The cartoon character's fall is delayed by the unerring horizontal direction of his movement. Correspondingly, the visual search used to follow his movement is horizontal and can remain horizontal for as long a period as the character's movement allows. His delayed fall allows for the reification of the infantile schema which allowed for movements of accommodation peculiar to preceding perceptions.

The second interpretation pertains to egocentric thought. In *The Child's Conception of Physical Causality*, Piaget (1930) explained that children see themselves as the center of causality in the universe. A well-known example cited by Piaget is the subject who says that the Arne River flows because men row their boats on it; when no oars are used in the water, the river does not flow. Another example is the child who believes that clouds move only when he does. In the case of the cartoon, it is significant that the character falls only when he *realizes* that he should. Hence, he actually determines the operation of physical law. His determination of gravity can be interpreted as a reification of residues of earlier structures which sanctioned egocentric thought.

[Example 3]. A magician pours the contents of a tall, thin beaker into a short, stout beaker. Then he pours the contents of the latter vessel into a tiny glass which is much shorter and narrower than each of the others. The act is performed before an audience of adults and adolescents.

It is conceivable that the performance reifies cognitive *residues* associated with conservation of liquids. Piaget (1952a) has proved that young children center their attention on either the height or width of a vessel to estimate the volume of liquid contained in it. If the child centers on height, he will say that the contents of a tall, thin beaker diminish when poured into a short, broad beaker. If he centers on width, he will hold the opposite viewpoint. Later, the child learns that the volume of liquid is conserved, regardless of the shape of its container. The magic act described above seems to present proof of the height-centered conception of conservation. By emptying the contents of the short, stout beaker into the tiny glass, the magician demonstrates that there "really" is more liquid in tall beakers than in short (and stout) ones. This is made clear by the profound difference in height between the first and third beakers in the pouring sequence. The narrowness of the third beaker is theoretically unimportant in the performance, since height alone is being promoted as the determinant of volume. The tininess of the third beaker helps add a sense of drama to the demonstration.

Two foreseeable challenges can be raised against the above interpretations. The first contests the ascription of residual bases to the fantasy expressions that have been described: Why should it be assumed that these examples of fantasy are related to cognitive residues? Isn't it possible that people simply enjoy seeing violations of physical laws? The challenge can be met on two grounds. First, the previous discussion of reification has pointed up the structural bases of reifying activity. When conflictive situations are purposely sought, it can be assumed either that curiosity is being exercised, in which case the range of reifying activity is being extended for existing structures, or that testing of newly evolved and still tentative constructs is occurring, as in *exploratory* fantasy. For the examples being considered, there is considerable evidence, especially from Piaget, that there were cognitive structures in the mind for which the specific fantasy presentations were *non*conflictive. If the argument for reification can be trusted, attributing fantasy to structures whose reifications were *least* conflictive, or in this case, *non*conflictive, with the specific fantasy expression is more justifiable than attributing fantasy to affinities for vaguely defined kinds of events. In other words, if structural homologies can be found in any expression of fantasy, these homologies probably reside in structures whose reifications can best be achieved in this fantasy expression. Invoking ambiguous likes and dislikes to explain the fantasy expression is both problematical from the standpoint of definition and nonproductive from the standpoint of further research.

The second response to the challenge is a continuation of the first. Granted that there were earlier structures whose reifications would have been nonconflictive with the above fantasy expressions, it can be shown that these earlier structures are still active in the cognitive framework. First, Piaget (1954) has observed residual behavior in which brief reversions are made to old strategies of action. Pinard and Laurendeau (1969) have commented that this behavior represents certain forms of "horizontal décalages," or developmental lags. The atavism represents a temporary return to former structures, although it occurs only briefly after modification of the structure. The second source of evidence for the existence of residues is Miller's (1971) research. Miller has found that 11-year-old conservers of weight are less confident about conservation than was originally thought. These conservers, Miller has found, can be induced to revert to their earlier appreciations of conservation. Whether similar research has been done for other concepts is not known by the author. A third argument can be made on a priori grounds: The retention of outmoded structures should prevent repetitions of the assimilative and accommodative processes that initially produced these same unsatisfactory schemata; the outmoded structures should stand as "reminders" of old, abandoned beliefs that should now be avoided. On the same account, the retention of the abandoned structures should help adults empathize with children by providing the adults with some recall of their own earlier beliefs.

The second challenge to the residual interpretations of the above fantasy expressions disputes the existence of any cognitive structures which could have served as foundations for the residues: The fantasy presentations only have a reifying effect on the so-called residues because they contradict the expectations of later structures; there is no reason, therefore, for supposing that the fantasy expressions directly reify cognitive residues. This is an important criticism. One is hard pressed to notice that if the same presentations were repeated for children who had not yet reached the object concept or conservation, the presentations would probably not be appreciated as reifying or nonreifying. No matter what kind of proof might be given for or against the permanence of objects, for example, the pre-object-concept child would still act as if he assumed object impermanence. The nonconserver would react identically to demonstrations supporting or contradicting conservation (Smedslund, 1961). According to the challenge being raised, it would be claimed, then, that there were in fact no earlier structures related to the phenomena that supposedly have a reifying value. It would be claimed, for instance, that it was not that pre-object-concept children had a definite belief in the impermanence of hidden objects; it was rather that these children simply had not yet achieved an understanding of the permanence of hidden objects. It would be claimed, to use another example, that it was not that preconservational children had a clearly defined belief in nonconservation; instead, they had not yet developed a conservational structure. The truth of these statements would preclude the existence of residues if these residues were derived from purely fictitious structures.

There appear to be two replies to the challenge. First, one can find residues derived from identifiable structures. Unlike the residues of Examples 1, 2, and 3, which were formed in the sensorimotor period, other residues can be traced to the preoperational and concrete operational stages. The existence of structures in the sensorimotor period is debatable, as is clear from the above challenge. But there can be no doubt of the presence of structures in the preoperational and, particularly, in the concrete operational stages; indeed, the latter stage is largely defined by the broad structuralization that occurs in it. Now two of the major belief "systems" characterizing the preoperational and concrete operational stages are animism and artificialism. These belief systems are discarded as maturation continues. Yet one has only to take a cursory look at fantasy literature, much of it written by adults, to see that it is replete with animistic and artificialistic references and themes. The reliance on animism and artificialism is, in fact, so common that specific examples of it are unnecessary here. It need only be said in sum that the expression of animism and artificialism in adult fantasy can be linked to the very same beliefs, which stood as prominent structures during childhood.[3]

The second reply to the challenge grants the possibility that some of the primary structures from which the residues originated may never have existed. This argument does not propose that the residues were formed ex nihilo. Instead, it proposes that at least some of the residues may have been formed ex post facto. In explaining absorption, Neisser (1962), for example, has cited the inability to decipher the contents of hidden-picture puzzles as representing constructs which later are absorbed into constructs that facilitate this deciphering. Remembering that absorption produces cognitive residues, it can be argued that residues could have been formed here even if they did not arise out of constructs per se, which prevented deciphering of the puzzle. After absorption, it became clear *that* the puzzle could not be deciphered before. *That* the puzzle could not be deciphered also entailed retention of *what* was inherent in the previous understanding, that is, the *non*understanding of the puzzle. Hence, a residue was formed despite the absence of a parent structure. This process may occur in the development of the object concept and conservation. In both of these cases, it is recalled *that* object concept and conservation were not grasped before. The former absence of these concepts was *what* became embodied in newly formed residues.

RESIDUAL MECHANISMS

The coexistence of residues and their structural successors necessarily causes conflict. The reifications of the two types of constructs radically clash, as must be the case if one construct replaced the other during modification. As Berlyne (1964) has stated:

> Conceptual conflict and its attendent disequilibrium will not emerge until some external stimulus pattern, verbal or nonverbal, or some thought process causes incompatible symbolic responses to be aroused in combination [p. 569].

Thus conflict between residues and new structures is induced by the reifications which originate in both of these kinds of constructs: The reifications of new structures contradict residual reifications; residual reifications contravene reifications of new structures. The conflict between residues and new structures is aggravated and prolonged by these contraventions, and particularly by the contravention of new structures.

---

[3] A counterargument may be raised that fantasy literature is written solely to satisfy children's interests and tastes; thus, the residual basis of the expression is minimal. The main reply to this challenge is that a large portion of fantasy literature is not written for children, but is directed towards adults and adolescents. Further, the absence of residual bases for fantasy literature would presuppose an intimate familiarity with children's tastes for all fantasy authors. Since this is not the case, it must be assumed that the authors partly draw from recollections of their own childhood beliefs.

How does the aggravation and prolongation occur? It has been repeated a number of times that structures are in some way predominant in the cognitive framework; an analogy between new structures and the ego is not entirely inappropriate in this regard. The predominance of new structures is attributable to their reifiability in the environment and to the less probable reifiability of cognitive residues. Because of the active nature of reification, that is, the way in which the environment is shaped to ensure it, structures that are most generally and easily reified are the ones that predominate. Modified structures that are not reified are somehow repressed, as for example in Schachtel's (1947) "childhood amnesia." When predominant structures are directly contradicted or not reified, they temporarily "lose their effectiveness" (to repeat Klein's phrase) and less actively ensure the repression of residues. Residues then have a renewed opportunity to become reified, since they, or their reifications, are repressed less stringently than before.

This process is responsible for the longevity of cognitive residues. Whenever a new structure is contradicted by residual reification, the structure "falters" momentarily and imposes less repression on residues. This allows the residues to seek more reification in the future. How does this occur?

The new structure "falters" only momentarily because of the relative speed with which residual reifications are recognized as fantastic. As soon as recognition is achieved, the information contained in residual reifications is, in effect, taken less seriously: what would have been considered as conflictive information is dismissed as fantastic; accordingly, compensatory or modificatory processes that would otherwise have been put into effect are not. Before recognition occurs, however, the reifiability of the new structure is put into question. Because of the reification tendency, an attempt is made to promote whichever cognitive schema seems best reified by the conflictive phenomenon. The least appropriate structure is the new one, while the most appropriate structure is, of course, the residue. Therefore, an attempt is begun to reinstate the residue into a position of prominence in the cognitive framework. This effort continues until it is recalled that the residue represents an old, abandoned belief that has been modified in development. At this point, the residue is repressed once again and the new structure regains its prominent status. In this course of events, the residue did manage to ascend briefly in the structural hierarchy. During the period of reinstatement, the residue indirectly gained reassurance of its validity. Thus encouraged, it now continues to seek reification in order to prove its complete reifiability and to facilitate its reinstatement as a predominant structure. This effort is futile because residual reification is recognized as such by predominant structures. But the effort will be continually repeated because recognition will always take some time during which predominant structures will "falter."

Owing to this theoretical process, the fading of residues is prevented. The mechanism is introduced because of evidence that residues remain "intact" in the mind throughout adulthood. The main source of evidence for the retention of residues is the enjoyment and production of residual fantasy by adults (see Examples 1, 2, and 3).

If residues remain intact in the mind, there are many such residues. The abundance of residues is largely due to the extensive structural modification that occurs in (early) development. But if many residues remain intact, this does not presuppose that each residue has its own separate reification in fantasy. Reifying each residue separately would raise serious practical problems concerning the time devoted to the expression of fantasy.

Instead of separate reification, residues can be expressed collectively. Individual residues can be subtly reified in a collective reification which is clearly atavistic and which can thus be easily identified as fantastic. The collective reification can be accomplished in the *style* or *method* of fantasy, as for example in the illogical progression of events in dreams or in the idiosyncratic style of artistic expression. Of course, structures are always reified collectively, since they are abstracted from and reapplied to a world of complex and interrelated events. Styles of behaving always reflect a wide array of cognitive structures, whether these structures are residual or not. But residual fantasy is unique in its reflection of former structures that have been modified. The chain of events in dreams, for example, exemplifies causality and associativity as these processes were conceived in infancy. Chagall's reliance on animism in his paintings may be linked to his infantile anismistic beliefs which, as Piaget (1929) has shown, are common in young children. The spatial and temporal incongruities which characterize many films seem imposingly similar to the same kinds of incongruities that have been outlined by Piaget.

The residual influences on fantasy style, then, seem as profound as the residual influences on fantasy content. The extent of this influence cannot be assessed here, mainly because the collective reification of residues is phenomenally distinct from the reifications of individual residues. As increased attention is devoted to the manifest content of dreams and fantasy (e.g., Eggan, 1961), the residual nature of fantasy style will hopefully be clarified.

PROPAGATION OF FANTASY

Perhaps the most significant implication of the theory of cognitive residues is the notion that fantasy breeds more fantasy. Regeneration of fantasy operates both in the individual and among individuals over generations.

Fantasy regeneration in the individual can be inferred from the positive feedback process which accounts for the longevity of residues. Residues remain "intact" because they are reified through fantasy. They are stimulated to seek further reification because predominant structures "falter" when confronted with residual fantasy. Because residues stay "intact," they continue to spawn fantasy.

Fantasy regeneration among individuals exemplifies Maruyama's (1963) "second cybernetics": man imposes delusional systems on the environment and responds to these systems; his imposition of the systems and his subsequent responses to them form a mutally causal relationship. The next example and the discussion that follows illustrate what is meant.

[Example 4]. A naive youngster sees a flying man depicted on a movie screen. The film of the flying man was made by a movie producer who wanted to reify a cognitive residue formed in his childhood when he saw a painting of an angel in flight. This picture was painted by an artist who wanted to reify a cognitive residue associated with some experience in his childhood.

Though the example is hypothetical, it outlines the regeneration of fantasy among individuals in successive generations.

Because neonates are unsophisticated judges of fantasy and "reality" (Taylor, 1971), their assimilations of fantastic and "realistic" percepts are indiscriminate. Consequently, their cognitive frameworks tend to internalize both kinds of percepts equally; cognitive structures which embody fantastic information are formed as readily as structures which embody realistic information. Later, when the distinction between fantasy and reality is learned, fantasy structures are modified and are subordinated to reality structures. Those fantasy structures which are displaced or absorbed become cognitive residues that are reified in fantasy.

The transmission of myths and legends exemplifies this process more concretely than Example 4. Stories told to children in successive generations are almost always fantastic. The stories are internalized by the children and are ostensively believed to be true. Cognitive structures which assume the contents of the tales are formed and are later modified when the fantasy–reality distinction is acquired. After this distinction is made, hearing the story is still enjoyable because residues induced by the story can be reified. Finally, it becomes satisfying to recount the tale to children of the next generation for the same purposes of reification.

An explicit assumption in the above paragraph is that myth transmission is perpetuated by the presence of residues induced by the myth itself; the residues could only have been formed if the myth was recognized as fantastic. Considerable debate has centered around the assessment of fantasy in myth and legends, especially as achieved by primitive peoples. Tyler (1871) and Lévy-Bruhl (1926) discounted the capacity in primitive peoples for fantasy recognition. More recently, Spiro (1964) and Spiro and D'Andrade (1958) have noted that youngsters naturally assume the truth of information transmitted by significant others; the learned information is proved empirically by the very fact that it comes from significant others and that behaviors based on the information are vital to the culture's projective institutions (Kardiner, 1945). On the other hand, Eliade (1963) has remarked that primitive peoples are capable of distinguishing "true" stories from "false" stories. And Hallowell (1955) has exclaimed that "Man's very existence depends on making *some* distinction between 'objective' and 'subjective' [p. 85]." All of the above authors agree that rationality is culturally relative. Their disagreement hinges on the importance ascribed to empirical evidence drawn from the "physical" world.

Without intending to resolve the disagreement, some conciliation can be derived from the research of Taylor (1971) and from the cross-cultural study of cognitive growth. Taylor found that young children have an increasing ability to distinguish fantasy from reality as they get older; the effect for age was found to be significant at the .01 level. Cross-cultural study has established the universal invariance of the developmental sequence (although differences have been found in the age of onset of certain stages and in the presence of late stages). In combining the two lines of research, the universal invariance of the developmental sequence should guarantee the generality of Taylor's findings cross-culturally. At the same stage in which American children discover the fictitious nature of Santa Claus, for instance, Senegalese children should grasp the fantasy attached to mythical characters specific to their culture.

Of course, the probability of fantasy recognition in different cultures should be a function of conflict induced by the fantasy. As has already been said, a central reaction to conflict is the attempt to prevent its recurrence. Recognizing conflict that has already been mitigated serves this end. With recognition, familiar conflict is immediately resolved, for recall of the conflict includes recall of the means used to resolve it. Conflict recognition, then, is tantamount to conflict prevention. Recognition especially comes into demand as conflict increases.

Now recognition becomes most effective when it becomes generalized, when it applies to whole classes or types of conflict rather than to specific situations which are dissonant. Recognition of a broad category such as fantasy, for instance, is highly adaptive, since fantasy is often responsible for vast portions of conflictive phenomena that irritate the child. Generally, this implies that the child is most likely to develop fantasy recognition as conflictive pressure imposed by fantasy is highest. Moreover, a plethora of fantasy-induced conflict should help the child recognize the "common denominator" of this conflict, which is fantasy itself. The capacity for fantasy recognition should certainly be determined to a large extent by the overall intellectual maturity of the child. But just as conflict catalyzes intellectual growth, so too should it spur the advent of fantasy recognition when fantasy itself imposes this conflict.

Once the child has been equipped with the ability to distinguish fantasy from reality, he can effect decisive modifications of fantasy-induced structures. Ironically, many of his modified structures continue to be expressed. On Christmas eve, for instance, when he does not see or hear Santa Claus climbing down the chimney, the young child finds Santa's absence inconsistent with his fantasy-based expectations. To mitigate the resultant dissonance, it becomes necessary to modify the belief in Santa Claus. The structure or structures which allowed for his existence are recognized as fantastic and may be assigned residual status. Still, seeing Santa Claus in a shopping mall next year, for instance, will be enjoyable. Years later it will be satisfying to expose children of the next generation to the jolly fat man in the red coat. Presenting Santa Claus to these children will cause them to believe in him, since he will have been introduced by a significant other (i.e., the parent). Transmission of the Santa Claus myth to children will also effect a wider population which accepts his authenticity. Much in line with Asch's (1955) findings, this expanded population will help socially substantiate the parent's (atavistic) belief in the real existence of Santa.

This process will repeat itself over and over again. Whether it will occur in analogous form in non-Western societies can at present only be a moot point. Nevertheless, it is hard to imagine that the germinating factor of contradiction and modification of "delusional" sets can be entirely prevented in any culture.

Technology

A central element in the propagation of fantasy is the technology it employs. It can be argued that the growth of technology has promoted an expansion and intensification of fantasy through the course of history.

Residual fantasy must be as "convincing" as possible for residues. Since cognitive structures, residual or not, are meant to be applied to events in the "real" world, reifications are most successful when they are high in verisimilitude: Predominant structures are best reified by actual events in the environment; residues are best reified by conditions which would confirm them in their predominant, premodified form.

Now technology can facilitate the verisimilitude of residual expressions of fantasy. The hypothetical "Santa Claus" residue, for example, can best be reified by dressing someone in a red suit, white beard, etc. By using these materials, the actor looks more like Santa than another actor who is undisguised. This is not to suggest that residues are best reified by reproducing those percepts which initially spawned preresidual structures. In the film of the flying man, for instance, this presentation was made, despite the fact that its initiating percept for the producer was a static picture. It can be reasoned safely, then, that the painted, representational percept of the flying man was abstracted into a more general construct which had an enactive reificational field. This process of abstraction to enactive events comprises a basic component of construct formation prior to the development of propositional logic in the formal operational stage.

The growth of technology has promoted an expansion and intensification of fantasy through the course of history, because technology has allowed for an increased verisimilitude of fantasy expression. As in the familiar case of the flying man, development of animation in cinema allowed for a more lifelike presentation of human flight than was possible in static paintings. The refinement of painting allowed for a more detailed and lifelike presentation than could have been achieved in primitive art.

How did the increased verisimilitude of fantasy expression cause an expansion and intensification of fantasy? First, a wider array of cognitive residues could be reified as technology developed. With more subtle means of artistic expression, more numerous residues could be expressed. Consequently, people in successive generations were exposed to more and more expressions of residual fantasy. More preresidual structures were formed as a result, and after becoming residues these structures required their own reification through fantasy. Second, preresidual structures were formed more readily, and thus more profusely, as the residual expressions of earlier generations became more lifelike. Recalling that the developing individual is first sensitive to enactive, then to ikonic, and finally to representational percepts (Bruner, Olver, & Greenfield, 1966), the very young child is prone to assimilate more straightforward presentations of fantasy than highly symbolic and abstract presentations. Also, because learning of the fantasy–reality distinction parallels development of sensitivity to ikonic and representational percepts, very young children are likely to assimilate a minimal number of symbolic presentations which are fantastic. What this implies is that children acquire more cognitive residues as fantasy expression becomes more lifelike. Furthermore, as technology facilitates more lifelike expression of fantasy, it also effects a broader range of communication. Thus more children acquire an increased number of residues as technology grows. When television is introduced in the home, for instance, more children are exposed to more lifelike expressions of fantasy at an earlier age than ever before. More cognitive residues are formed as a result and the children's subsequent motivation for residual fantasy becomes more intense than it was for children of previous generations.

Up to a point, the system described here is self-serving. Dynamic equilibrium is maintained so long as individuals can reify their residues as effectively as those whose reifications initially caused formation of the residues. For example, the movie producer who had access to filming equipment and a studio could create a very lifelike reification of his "human flight" residue. The youngster who formed a flight residue after seeing the producer's movie could only reify this residue as effectively as the producer if he also had access to these resources.

But it is unlikely that the youngster ever gains access to these means, for only a limited segment of the population goes into film production. Consequently, he continues to enjoy watching the movie or he seeks other means of residual reification. Repeatedly watching the movie is made impossible by the contingencies of school and, later, of the vocation. So other means have to be pursued. If these means are nontechnological, their usage entails a disruption of the continuum between technology and fantasy expression.

Such a disruption may exist in America today. It is possible that the grown-up children of the 1950's are taking hallucinogenic drugs more frequently and commonly than people of earlier generations, because they experienced television in the home while their predecessors did not. TV programs conveyed fantasy information in a lifelike and easily accessible way. Abundant preresidual structures were consequently formed in the minds of young TV viewers. Following modification of the structures, an intense "need" for reifying fantasy existed. Without access to technological means of reification which were as effective as television, and prevented from watching television continuously (excluding a boredom factor), grown-up TV watchers sought some cathartic, nontechnological means of fantasy expression. One of the only recourses available may have been in hallucinogenic drugs.

This is not to dismiss a plainly emotional factor in drug use. What is urged here is that behavioral scientists acknowledge that drug-induced hallucinations may be as epistemologically rooted as drug-induced "highs" may have emotional roots and functions.

EDUCATION

Not all grown-up TV viewers take drugs and only a small minority are drug addicts. Abstinence from drug use may be attributable to certain educational practices which help to limit the formation of cognitive residues in children.

Preresidual structures are formed when new information is assimilated into the cognitive framework. The structures are modified when they fail to be reified or when their reifications are recognized as being specific to the context of fantasy.

It is possible that parents help reduce the "need" for fantasy in their children by limiting the number of preresidual constructs that are formed. They may do this in two ways. First, they may restrict the child's exposure to fantasy expression. Second, they may clearly indicate contextual differences between fantasy and reality for the child from the time that he is very young. The first strategy is hardly practicable, since it demands isolation of the child from other children who are exposed to fantasy and removal of the child from the almost ubiquitous media. In addition, the tactic retards acquisition of the fantasy–reality distinction that becomes vital when the child grows up and almost inevitably encounters at least some fantasy. The second is more likely to succeed. By providing obvious cues meant to show the difference between fantasy and reality, parents speed the understanding of this distinction in their children. From the time that the understanding takes root, children are less prone to construct structures based on fantastic percepts. Consequently, fewer residues are formed in the children's minds and less residual reification becomes necessary in their later lives.

To illustrate this educational technique, again consider the boy who sees the movie of the flying man. Assuming that the percept of the flying man makes a profound impression on the youngster, he will verge on incorporating it into his cognitive framework. Being naive, he probably will not have a specific construct which denies human flight. He will be caught in a dilemma between his understanding of human locomotion as it is seen in the movie theater and as it is seen and experienced in the "operational" world. His dilemma may be resolved by his parents' influence. They may provide verbal and nonverbal attitudinal cues about the film. "It's really amazing what people can do in *make believe!*" they may say to him. Their skepticism about human flight may enable the boy to reach an intuitive conclusion about the contextual differences between the theater and the outside world. He may then form a construct which denies human flight instead of forming a construct which allows for it. Since the latter construct would have eventually become a residue, its absence makes for one less residue that later seeks reification in fantasy.

If parental guidance can, in fact, effect isolated decisions about reality and fantasy, then there might also be ways to have this guidance instill a more general grasp of the fantasy–reality distinction at an early age. Two strategies would seem plausible here. First, parents could indicate the fantasy–reality distinction in a wide variety of experience. Second, parents could expose their young to obviously fantastic presentations to exaggerate the contrast between what is real and what is fantastic. Instead of giving a child a doll which walks, talks, cries, and wets her pants, for instance, it might be more desirable to give a simple rag doll which is much less like a human infant. Possession of the rag doll would perhaps make it easier for the child to determine the nature of dolls versus human babies. Incorrect structures would then be avoided and would subsequently not become residues. In similar fashion, clearly fictitious television programs and movies might be prescribed over more subtle presentations. The underlying objective here would be the simplification of fantasy recognition. Mastery of simple fantasy recognition could prove helpful in assessing more subtle expressions in the future, which would help further to limit the formation of residues.

CONCLUSION

Most of the ideas presented here have been meant to suggest possibilities. The concepts, though based directly and indirectly on some empirical evidence, require more proof. Residual bases of fantasy will have to be tested systematically by tracing the expression of (modified) childhood structures in fantasy. Conversely, fantasy expressions will have to be assessed for their linkages to earlier structures. Functions between fantasy and age, education, and access to technological resources will have to be substantiated. The correlation between drug use and the above variables will have to be established less speculatively than it has been here. Laboratory and field research, cross-cultural, and longitudinal studies will have to be implemented to support or weaken the ideas that have been presented.

Despite empirical shortcomings, however, the theory of cognitive residues seems to offer two seminal ideas. First, it diverts attention from the study of fantasy as it has been couched in psychoanalysis. Interpreting fantasy as an epistemologically exploratory and residual exercise has permitted the investigation of fantasy expression in its manifest as well as in its symbolic form. The manifest content of fantasy is puzzling in its own right as a cognitive operation, so that study of fantasy on the purely symbolic level now seems a bit premature, if not overly ambitious. This is not to say that the symbolic approach should be repudiated. Nor should it just be fitted into a compromise with manifest content. Symbolic as well as manifest functions should be recognized as having a joint place in fantasy expression. Wish fulfillment, exploratory, and residual processes should be seen as working jointly: Wish fulfillment and exploratory processes should be able to co-occur, as in projection into unfamiliar situations of a sexual nature; wish fulfillment and residual processes should be able to co-occur, as in adult Oedipal fantasies: exploratory and residual processes should be able to co-occur as in fantasies of residues formed ex post facto which are then seen in the new perspective of later structures. If the different processes do not occur simultaneously, they should be able to occur in series in general fantasy expressions such as dreams.

The second important idea rests in the theory's implicit commentary on irrationality. Any behavior performed by a "rational" person is considered irrational if it violates that person's understanding. Classically, the irrational behavior is explained as affect. The present theory, however, poses another explanation: The person's irrational behavior is actually rational by old standards or rational by standards that are being tried on for size. The latter case becomes hard to measure, since tested systems can be rejected before they ever pass the testing stage. The former case can be easily assessed by comparing the individual's cognitive history and his irrational behavior. Finding a significant connection between residues and wide varieties of irrational behavior and fantasy for a large sample would help to substantiate the belief that man's behavior is governed by rational rather than by irrational forces.

REFERENCES

ABELSON, R. Modes of resolution of belief dilemmas. Paper presented at the meeting of the Western Psychological Association, Monterey, California, 1958.

ASCH, S. Opinions and social pressure. *Scientific American,* 1955, **193**, 31–35.

BERLYNE, D. Uncertainty and epistemic curiosity. *British Journal of Psychology,* 1962, **53**, 27–34.

BERLYNE, D. Toward a theory of epistemic behavior: Conceptual conflict and epistemic curiosity. In R. Harper, C. Anderson, C. Christenson, & S. Hunka (Eds.), *The cognitive processes: Readings.* Englewood Cliffs, N. J.: Prentice-Hall, 1964.

BOWER, T. The object in the world of the infant. *Scientific American,* 1971, **225**, 30–38.

BRUNER, J., OLVER, R., & GREENFIELD, P. *Studies in cognitive growth.* New York: Wiley, 1966.

BURTON, R., & WHITING, J. The absent father and cross-sex identity. *Merrill-Palmer Quarterly,* 1961, 7, 85–95.

CHARLESWORTH, W. The role of surprise in cognitive development. In D. Elkind & J. Flavell (Eds.), *Studies in cognitive development.* New York: Oxford University Press, 1969.

DEVEREUX, G., & LABARRE, W. Art and mythology. In B. Kaplan (Ed.), *Studying personality cross-culturally.* New York: Harper & Row, 1961.

EGGAN, D. Dream analysis. In B. Kaplan (Ed.), *Studying personality cross-culturally.* New York: Harper & Row, 1961.

EHRENZWEIG, A. A new psychoanalytical approach to aesthetics. *British Journal of Aesthetics,* 1962, 2, 301–317.

ELIADE, M. *Myth and reality.* New York: Harper & Row, 1963.

FESTINGER, L. *A theory of cognitive dissonance.* Evanston, Ill.: Row, Peterson, 1957.

FESTINGER, L. The motivating effects of cognitive dissonance. In R. Harper, C. Anderson, C. Christenson, & S. Hunka (Eds.), *The cognitive processes: Readings.* Englewood Cliffs, N. J.: Prentice-Hall, 1964.

FLAVELL, J. Stage-related properties of cognitive development. *Cognitive Psychology,* 1971, 2, 421–453.

FLAVELL, J., & WOHLWILL, J. Formal and functional aspects of cognitive development. In D. Elkind & J. Flavell (Eds.), *Studies in cognitive development.* New York: Oxford University Press, 1969.

FREUD, S. Moses des Michelangelo. *Imago,* 1914, III.

FREUD, S. *Leonardo da Vinci.* New York: Moffat, Yard, 1916.

HALLOWELL, A. *Culture and experience.* Philadelphia: University of Pennsylvania Press, 1955.

HEBB, D. The mammal and his environment. *American Journal of Psychiatry,* 1955, **111**, 826–831.

HELD, R. Plasticity in sensory-motor systems. *Scientific American,* 1965, **213**, 84–94.

HOLLOWAY, R. Culture: A human domain. *Current Anthropology,* 1969, **10**, 395–412.

HUNT, J. *Intelligence and experience.* New York: Ronald Press, 1961.

JACOBSON, E. Response to a sudden unexpected stimulus. *Journal of Experimental Psychology,* 1926, 9, 19–25.

KARDINER, A. *The psychological frontiers of society.* New York: Columbia University Press, 1945.

KELLY, G. *A theory of personality.* New York: Norton, 1963.

KLEIN, G. Cognitive control and motivation. In G. Lindzey (Ed.), *Assessment of human motives.* New York: Rinehart, 1958.

LANDIS, C., & HUNT, W. *The startle pattern.* New York: Farran & Rinehart, 1939.

LANE, M. (Ed.). *Structuralism: A reader.* London: Jonathan Cape, 1970.

LÉVY-BRUHL, L. *How natives think.* London: G. Allen, 1926.

MARUYAMA, M. The second cybernetics: Deviation and amplification, mutual causal processes. *American Scientist,* 1963, **51**, 164–179.

MEAD, G. H. *Mind, self, and society from the standpoint of a social behaviorist.* Chicago: University of Chicago Press, 1934.

MILLER, G., GALANTER, E., & PRIBRAM, K. *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston, 1960.

MILLER, S. Contradiction, surprise, and cognitive change: The effects of disconfirmation of beliefs on conservers and non-conservers. Unpublished doctoral dissertation, University of Minnesota, 1971.

Neisser, U. Cultural and cognitive discontinuity. In The Anthropological Society of Washington, *Anthropology and human behavior.* Washington, D. C.: Gaus, 1962.

OLDFIELD, R. Memory mechanisms and the theory of schemata. In R. Harper, C. Anderson, C. Christenson, & S. Hunka (Eds.), *The cognitive processes: Readings.* Englewood Cliffs, N. J.: Prentice-Hall, 1964.

PARSONS, T. Social structure and the development of personality. In B. Kaplan (Ed.), *Studying personality cross-culturally.* New York: Harper & Row, 1961.

PIAGET, J. *The child's conception of the world.* London: K. Paul, Trench, Trubner, 1929.

PIAGET, J. *The child's conception of physical causality.* London: Kegan Paul, 1930.

PIAGET, J. *The child's conception of number.* London: Routledge & Paul, 1952. (a)

PIAGET, J. *The origins of intelligence in children.* New York: International Universities Press, 1952. (b)

PIAGET, J. *The construction of reality in the child.* New York: Basic Books, 1954.

PIAGET, J. Logique et equilibre dans les comportements du sujet. In L. Apostel, B. Mandelbrot, & J. Piaget (Eds.), *Études d'epistémologie génétique.* II. *Logique et equilibre.* Paris: Presses Universitaires de France, 1957.

PIAGET, J. *Structuralism.* New York: Basic Books, 1970.

PINARD, A., & LAURENDEAU, M. "Stage" in Piaget's cognitive-developmental theory: Exegesis of a concept. In D. Elkind & J. Flavell (Eds.), *Studies in cognitive development.* New York: Oxford University Press, 1969.

PRENTICE, W. *Some cognitive aspects of motivation.* In R. Harper, C, Anderson, C. Christenson, & S. Hunka (Eds.), *The cognitive processes: Readings.* Englewood Cliffs, N. J.: Prentice-Hall, 1964.

RAPAPORT, D. The theory of ego autonomy: A generalization. *Bulletin of the Menninger Clinic,* 1958, **22**, 13–35.

SANFORD, N. The dynamics of identification. *Psychological Review,* 1955, **62**, 106–117.

SCHACHTEL, E. On memory and childhood amnesia. *Psychiatry,* 1947, **10**, 1–26.

SCHACHTER, S. *The psychology of affiliation.* Stanford: Stanford University Press, 1959.

SKAGGS, E. Changes in pulse, breathing, and steadiness under conditions of startledness and excited expectancy. *Journal of Comparative Psychology,* 1926, **9**, 303–317.

SMEDSLUND, J. The acquisition of conservation of substance and weight in children. II. External reinforcement of conservation of substance and weight and of the operations of addition and subtraction. *Scandinavian Journal of Psychology,* 1961, **2**, 71–84.

SPIRO, M. Religion and the irrational. Paper presented at the meeting of the American Ethnological Society, Pittsburgh, 1964.

SPIRO, M., & D'ANDRADE, R. A cross-cultural study of some supernatural beliefs. *American Anthropologist,* 1958, **60**, 456–466.

STRAUSS, H. Das Zusammenschrenken. *Journal für Psychologie und Neurologie,* 1929, **39**, 111–231.

TAYLOR, B. The ability of three-, four-, and five-year-old children to distinguish fantasy from reality. Unpublished doctoral dissertation, Brigham Young University, 1971.

TROLAND, L. *The principles of psychophysiology.* Vol. 3. *Cerebration and action.* New York: Van Nostrand, 1932.

TYLER, E. *Primitive culture.* London: John Murray, 1871.

WAELDER, R. *Psychoanalytic avenues to art. Psychoanalytical epitomes.* New York: International Universities Press, 1965.

WALLACH, H. Memory effects in perception. Paper presented at the meeting of the International Congress of Psychology, Montreal, 1956.

ZANGWILL, O. A study of the significance of attitude recognition. *British Journal of Psychology,* 1937, **28**, 12–17.

(Received January 28, 1972;
revision received June 19, 1972)

Psychological Review
1972, Vol. 79, No. 6, 487–509

RHYTHMIC (HIERARCHICAL) VERSUS SERIAL STRUCTURE IN SPEECH AND OTHER BEHAVIOR[1]

JAMES G. MARTIN[2]

University of Maryland

Lashley regarded rhythmic action as a central concept in the problem of serial order in behavior, and implied that it was a natural link between the perception and production of connected speech. Rhythm means relative timing between adjacent and nonadjacent elements in a behavior sequence (hence nonsequential dependencies), as opposed to temporal ordering (concatenation) of elements, with implications for motor and perceptual behavior in real time. Two simple rules for generating rhythmic (hierarchically organized) behavior sequences are proposed and applied to speech, music, and other meaningful sound patterns. The temporal constraints on rhythmic sequences lead to several hypotheses concerning their efficient perception and production.

Just over 20 years ago, Lashley (1951) published his influential paper on the problem of serial order in behavior, emphasizing speech and discussing a number of topics he believed to be highly interrelated, including associative-chain versus hierachical conceptions of behavior, determining tendencies, the generality of syntax, central control of motor patterns, rhythmic action, and others. All but one of the topics mentioned have been widely discussed, the first in particular (e.g., Adams, 1968; Chomsky, 1959; Greenwald, 1970; Halwes & Jenkins, 1971; Keele, 1968; MacNeilage, 1970; Miller, Galanter, & Pribram, 1960; Milner, 1970; Neisser, 1967; Wickelgren, 1969). The last of these topics, Lashley's views on rhythmic action, has not been considered in detail by anyone with the exception of Lenneberg (1967). But Lashley (1951) strongly implied that rhythmic action and hierachical motor organization were highly related concepts and even went on to hint that rhythmic action might be the natural link between the perception and production of connected speech. This latter theme, which is potentially one of the most interesting in his entire paper, has attracted little attention. To be sure, Lashley was not quite clear on the point, but another reason for its neglect could be the several misconceptions that are commonly held about rhythmic action. One of these is that rhythm implies only periodic, repetitive behavior, like walking or breathing, and hence is too simple a concept to be of interest for speech behavior. Another quite different misconception is that rhythm necessarily entails *strict* temporal regularities, usually along with complex, aesthetically motivated elaborations, as in music, and hence is too esoteric to apply to natural speech sounds except by trivial analogy. Lashley's (1951) views were not limited in either of these ways. By making clear what rhythmic action is and by showing how it has consequences for theories of motor and perceptual behavior in real time, this paper suggests that rhythmic action is one of the central concepts in the schema Lashley presented.

Rhythm, which means temporal patterning, is a concept based on motor functioning. Natural movement sequences as produced by the motor system are determined by a variety of factors including temporal constraints,

---

[1] This article was supported by Grant MH 16726 from the National Institute of Mental Health, United States Public Health Service, and by a grant from the Bio-Medical Sciences Support Committee, University of Maryland, to the Center for Language and Cognition.

[2] The author expresses thanks to Charles N. Cofer and to an anonymous referee for extensive comments on an earlier version. Nancy Hurlbut, Cathy Klemic, Dianne Reamy, Judith Sprei, and Judith Zachai assisted in the data collection and analysis.

Reqests for reprints should be sent to James G. Martin, Department of Psychology, University of Maryland, College Park, Maryland 20742.

some of which are based on the dynamics of inertia, etc. These constraints necessarily are reflected in the organization of sounds produced by human movements, which include speech and music. In the case of speech these constraints influence the segmental details of the acoustic signal, but they should be expected also to affect at least some aspects of the morphology and syntax of any language at the level of syllable strings since they play a role in determining what can be naturally and easily spoken.

The constraint on speech sounds, or on any other real-time sequence of behavioral elements, that is directly implied by the concept of rhythm is *relative timing,* which means that the locus of each (sound) element along the time dimension is determined relative to the locus of *all* other elements in the sequence, adjacent *and* nonadjacent. This is to say that sequences of sounds, speech or otherwise, that are rhythmic will possess hierachical organization, that is, a coherent internal structure, at the *sound* level. The alternative to a rhythmic sequence, which is less restrictive, is that the loci of the elements in the sequence are only successive (concatenated) in time. Sound sequences like these cannot have a structured, internal organization, although entities that they might represent can, of course, have such an organization, albeit one that is relatively highly abstract.

These alternatives, relative timing (rhythmic patterning) as opposed to concatenation, are related to quite distinct theoretical points of view concerning real-time behavior. It is more obvious that speech and other temporally unfolding behavior sequences proceed left to right than that they might possess internal structure. Hence it is relatively easy to adopt a concatenative (associative) view about these sequences, which will be reinforced by the inevitable finding that there are dependencies between adjacent elements in the sequence. Local (adjacent) dependencies, however, rarely reveal the role that these elements play in the more wholistic, hierarchically organized structures of which they are a part. The alternative to the associative view is to begin by carving nature at the joints, so to speak, by considering first the wholistic unit which *has* the internal structure and then by analyzing the dependency relationships between all elements in terms of the unit. Witness Chomsky's (1957) revolution in linguistics, which depended in part on his choice of the sentence rather than some other unit as the object for study. Wholistic real-time sound units (i.e., rhythmic patterns) also have a hierachical internal structure, as is demonstrated later in the paper.

Relative timing as opposed to concatenation also has logically distinct implications for production and perception models. The production of rhythmic sounds entails that the temporal locus, hence duration, of each sound element is related to each other locus in the resulting pattern. This requirement places a heavier constraint on timing mechanisms than does the concatenative alternative, which requires only that the resulting sounds be produced in the correct temporal order. Furthermore, the temporal patterning relations between nonadjacent sounds, hence movements, which rhythm implies seem to call for present timing, hence a predominantly central rather than peripheral control of the timing mechanisms.

In the case of perception, rhythmic constraints on production entail that the sound inputs during perception will be temporally *patterned.* Patterned speech sounds could be redundant with respect to linguistic message elements to a far greater extent than sounds that are only concatenated. Furthermore, since rhythmically patterned sounds have a time trajectory that can be tracked without continuous monitoring, perception of initial elements in a pattern allows later elements to be *anticipated* in real time. Suppose that some elements in a sequence are more informative than others. If these informative elements are nonadjacent and temporally predictable, then certain efficient perceptual strategies (e.g., attention cycling between input and processing) might be facilitated. Perception of concatenated sounds, on the other hand, would seem to require continuous attention. More generally, perceptual requirements should be expected to differ for sounds that can be expected at times T_1, T_K, *and* T_N, as opposed to sounds expected at unspecified times. These and other logical implications of relative timing for production and perception should be translatable into empirical consequences.

There is, however, no grammar and no model of speech perception or production that incorporates the constraints imposed by relative timing to speech sequences extending beyond a few phonemes; in all present grammars and psychological models, temporal relationships existing between elements are represented as concatenations only.[3] These concatenations will serve most grammars adequately, and even many psychological models, for example, models of sentence comprehension. Since the relationships of interest in such models are abstract, and since subjects in experiments testing these models will be provided either with inputs spoken by humans or inputs they can read, problems of accounting for the processing of the speech sounds themselves are bypassed and questions concerning any relations between the abstract entities and the sounds do not arise. But if relative timing is the true situation in speech, it has consequences for any theory or model concerned with connected speech in real time. Up to the present time the only such theories or models have been those in the theoretically and practically important areas of automatic speech synthesis and automatic speech recognition. The current status of these areas rather accurately reflects present understanding of speech production and perception, and in each of them relative timing could be an important consideration. A machine could produce concatenated speech, which in the present view would be temporally distorted but which would be fairly intelligible to humans nevertheless—given the high redundancy in speech (Ladefoged, 1967b) and the perceptual strategies available to the native listener. However, a general-purpose machine routine for recognizing natural connected speech that ignored temporal relationships would be inefficient at best and might not work at all if much of the more critical, or more reliable, information in the acoustic signal were systematically related to the temporal dimension of the signal.

In the next section, the meaning of rhythmic action is made explicit by means of two formal descriptive rules (a grammar) that are proposed for describing the timing of elements in a natural temporal pattern. These rules generate element sequences having a hierarchical structure that are here called rhythmic patterns. In the section following, these rules for rhythmic structure are applied to the patterns of speech, and then some implications for the phonetic reality of these patterns are considered. If speech has the rhythmic structure that the rules describe, then constraints are placed on the possible mechanisms that might be postulated for efficient speech perception and production. Several hypotheses for a speaker–listener model are presented, together with bits of evidence that might be relevant to each. The hypotheses are related to the general notion that speaking and listening are dynamically coupled rhythmic activities, such that linguistic information is encoded rhythmically into the signal by the speaker and decoded out of it on that basis by the listener. Finally, some speculations concerning the implications of rhythm for other research areas are offered. The emphasis in this paper is upon speech, but music and other meaningful temporal patterns are considered also, on the theory that they have a common base and are rooted in the motor and perceptual capabilities of the human organism.

RHYTHMIC PATTERNS

Two rules for the formal description of simple natural rhythmic patterns are proposed here. They are intended to relate a variety of intuitions and other facts about rhythmic structure, some of which are mentioned during the discussion.

Rhythmic patterns are defined as event sequences in which some events (elements) are marked from others, for example, loud (or high) sound (or left foot) versus soft

---

[3] Nor is there a theory of auditory temporal pattern perception which uses relative timing. In fact, except for Garner's work (e.g., Garner, 1970; Garner & Gottwald, 1968; Preusser et al., 1971; Royer & Garner, 1966, 1970) there has been little research on auditory temporal pattern perception since about the time of Woodrow (1909).

(or low) sound (or right foot); call the marked elements "accents." The accents recur with some regularity, regardless of tempo (fast, slow) or tempo *changes* within the pattern (accelerate, retard). Together an accented and unaccented element constitute the simplest rhythmic pattern; more complex patterns consist of patterns within patterns; that is, they are hierarchically organized (Martin, 1969b, 1970b).

Natural rhythmic patterns can take a variety of surface forms, but nevertheless have a simple underlying structure that can be characterized by one or the other of two obligatory rules. These rules express the relation between (*a*) relative accent level on the elements in a sequence and (*b*) relative timing of the elements.

The first rule is called the accent rule and applies to more or less continuing pattern sequences, the simplest case being the two elements represented by filled dots in Figure 1a. The numerals *under* the tree refer to relative accent level on the two elements (level one highest), showing that the accent naturally falls on the first element of a duple. In the four-element sequence represented by the four equidistant dots in Figure 1b, the strongest accent is on the first element, but the second strongest is on the third element, indicating the generally alternating character of accent level.

Now notice that left branches are labeled zero, right branches are labeled one. These represent binary numbers that can be obtained by reading up or down the tree and are used to determine either the accent level of a given element in the sequence or its serial position (not labeled) in the sequence. To compute accent level, read up the tree, convert the binary number to decimal number, and add 1. Recall that binary number 11 equals decimal number 3, 10 equals decimal number 2, 01 equals decimal number 1, and so on. Hence, accent level on the last element in Figure 1b, for example, is 3 + 1 = 4. Similarly, to obtain serial position, read down the tree, convert binary to decimal, and add 1. This reciprocal relation between accent level and accent location (timing) in the sequence is the accent rule and is general for any number of nodes.

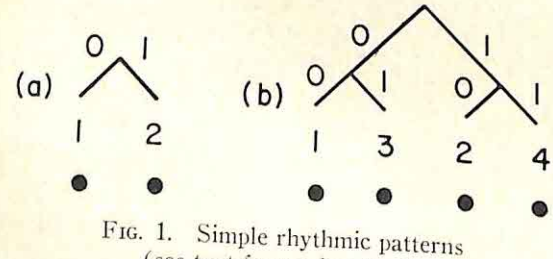


FIG. 1. Simple rhythmic patterns (see text for explanation).

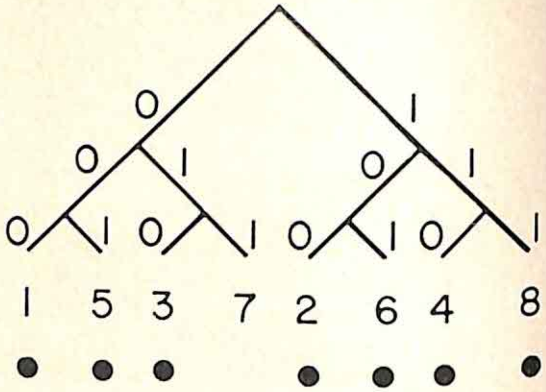


FIG. 2. Rhythmic tree with one null branch (see text for explanation).

The three-node-level tree in Figure 2 has a null branch, that is, the fourth element is missing, but with no effect on the remaining elements since accent level is determined not by serial position but by location along the time line, whether or not the element in question is actually realized. This applies not only to staccato elements, which are separated by silent intervals as in Figure 2 but also to more or less continuous sounds such as legato musical phrases or strings of syllables, in which case the temporal patterning would refer to the onset of each musical note or syllabic vowel.[4] Notice further in Figure 2 that the third element is necessarily lengthened if the sound sequence is continuous. Thus the duration of an element in continuous sounds depends upon rhythmic context and is rhythmically independent of its accent level, from the point of view of relative timing. Accent level correlates with duration, but for other reasons, as discussed later.

Notice in Figure 2 that the elements in first and fifth position are those most

---

[4] As is seen later, the stipulation that lower-level right branches can be null, that is, are somewhat optional, permits great flexibility in application of the rule.

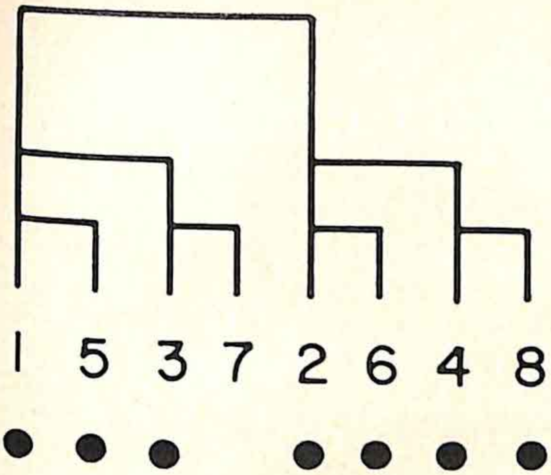


FIG. 3. Another way of drawing the rhythmic tree seen in Figure 2.

strongly accented. To demonstrate to yourself that these accent levels are the "natural" ones, tap an eight-element sequence on your desk, repeatedly, as fast as you can, accenting the first and the fifth elements. Then try the same thing but accent the first and the sixth elements. This pattern is more difficult than the first one, but not merely because accenting elements one and five degenerates the eight-pattern into two four-patterns. To demonstrate this, tap again, comparing the pattern having accents on elements one and six with another having accents on elements one, three, and five. The latter does not degenerate into equal-sized sequences, and it requires programming an additional accent, but it is generally easier nonetheless. Related to the view that the accent rule describes natural patterns is an observation by Woodrow (1951). He noted that listeners presented with continuous sequences of evenly spaced identical sounds often reported subjective grouping into fours, with stronger subjective accent on the first, and weaker subjective accent on the third, element of the group.

Figure 3 is exactly equivalent to Figure 2 and shows more clearly the relation between timing and accent levels. Time intervals are subdivided to the right via lower nodes but without affecting the temporal and accentual relations between elements springing from higher nodes. In this sense these elements are hierarchically rather than serially dependent. To the extent that right branches on lower nodes are optional, this tree maps onto a variety of terminal strings, but they will all have the same underlying structure; this is inherent in the meaning of rhythmic structure.

To illustrate its use, the accent rule in tree notation is applied below to some simple musical phrases, with appropriate musical notation added accordingly (pitch notation omitted). For the nonmusical reader it may be helpful to point out that nothing of relevance is missed if he prefers to read the tree only and to ignore the musical notation. The musical reader may appreciate that the musical notation is a shorthand for the tree diagram but, more important, that the tree diagram makes explicit the accent patterns usually implied by musical notation and tacitly known though never taught. Tree diagrams are as little used for learning music as for learning to read.[5]

In Figure 4, the vertical bars are boundaries between measures giving a two-measure phrase. Notes connected together by horizontal bars are subdivisions of one beat, giving two beats to a measure. The word *Old* falls on the first half of the first beat and receives strongest accent, the word *had* receives second strongest accent, and so forth.

In Figure 5, the first beat is further sub-

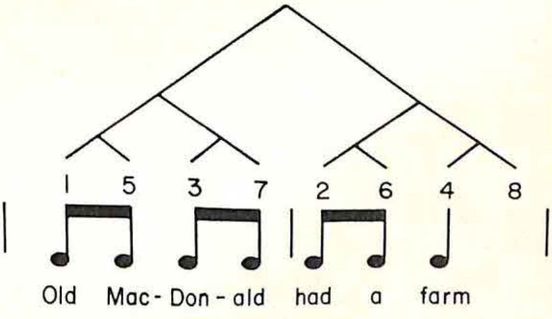


FIG. 4. The first phrase of a familiar tune. (Accent levels generated by the rhythmic tree can be seen above the notes.)

---

[5] That is, the rhythm rule as stated here makes explicit what the musician (or schoolchild) knows tacitly in the same sense in which a linguistic rule makes explicit what the speaker of the language knows tacitly. Rhythm appears to be taken so much for granted in music training that there is only one book on rhythmic theory although there are many on melody, harmony, and counterpoint (Cooper & Meyer, 1960). The latter are culturally determined to a far greater extent than is rhythm.

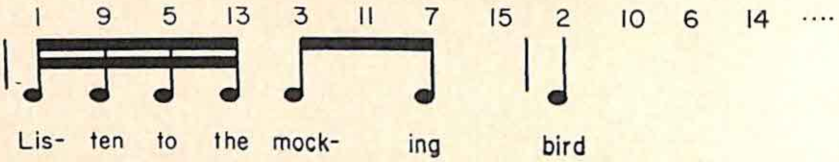


FIG. 5. Another familiar musical phrase. (Accent levels are those generated by a tree with four node levels.)

divided (each bar connecting notes halves their durations), requiring another node level in the tree but not a change in the relations between the three (higher level) notes on the beat. The tree has been omitted to save space, but it should be clear that the sequence of accent levels above the notes in Figure 5 is that generated by the accent rule for a four-node-level tree. The number of node levels is strictly determined by the element having shortest relative duration. As Figure 5 shows, adding a node level generally results in increased numbers of null branches.

Subdivisions of an interval can occur in other ways as well, as in the incomplete phrase in Figure 6. Examples like these show that the pattern of notes giving a tune its own rhythmic flavor can vary but nonetheless will map onto trees defined by the obligatory accent rule.[6]

Whereas the accent rule applies more or less to repeating patterns or to nonfinal musical phrases, the second rule, called the terminal rule, applies to nonrepeating or terminal sequences like musical cadences, as well as to a variety of speech units. Stated formally, it is: First, apply the accent rule to the tree to obtain the sequence of accents, then reverse the positions of accent levels one and two. The consequence of this rule, as is apparent later, is that it allows primary accent to appear not only on the initial element of the sequence but anywhere in the pattern. It is descriptively far more flexible than the accent rule.

The detailed mechanics for use of the terminal rule are shown later. For the moment it is convenient to consider an example of the results of the rule. Figure 7 shows the last two measures of the *Old MacDonald* song. Notice that the sequence of accent levels (shown above the notes) is exactly like the output of the accent rule for the three-node-level tree seen in Figure 4 except that accent levels one and two are reversed. In this context the effect is that the phrase then ends on a strong accent, which is typical of most simple musical cadences. In the next section, the terminal rule is shown to apply also to the rhythmic patterning of speech.

RHYTHMIC STRUCTURE OF SPEECH

"Rhythm" here as in conventional usage refers to the pattern of accents/stresses on

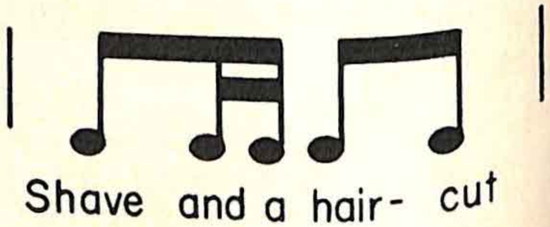


FIG. 6. Part of a familiar cadence.

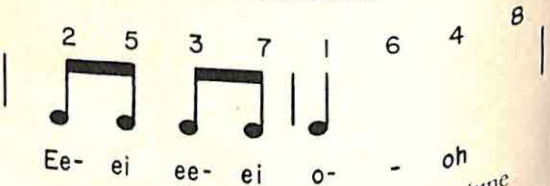


FIG. 7. The last phrase of a familiar tune (see text for explanation).

---

[6] Of course musical rhythms quickly depart from or elaborate upon simple rules such as these, but notice that syncopation and other interesting devices are effective only when played against the simpler underlying patterns, real or implied. More to the point here, complex rhythms are not required for ordinary speech and probably most poetry. Thus these rules have not been stated in music theory because they are too simple to be of interest. They have not, however, been stated in studies of linguistics or prosody either, to the best of the author's knowledge, although they are of theoretical importance in these areas. However, the importance of relative timing, etc., in verse has been known (or argued) for a long time (see, e.g., Lanier [1880]).

Incidentally, any reader still waiting for the resolution of the incomplete phrase in Fig. 6 is probably aware that the strong accent beginning the next measure is not actually sounded but only implied (and felt), a conventional rhythmic device.

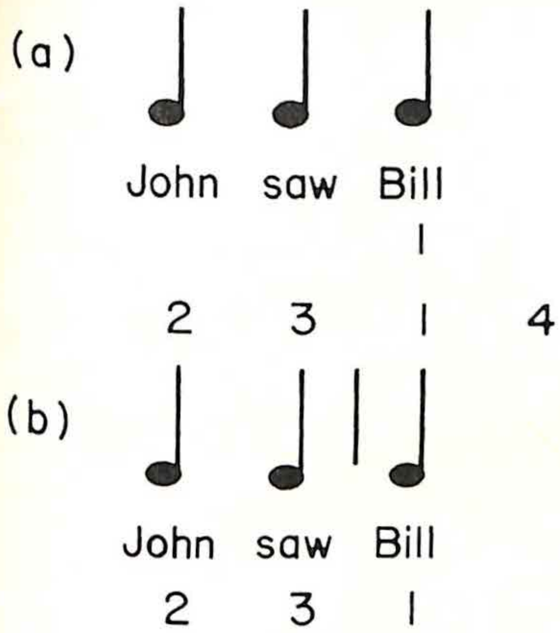


FIG. 8. Applying the terminal rule to a simple sentence (see text for explanation).

a string of syllables. The rhythmic pattern as defined here is assumed ordinarily to consist of up to seven syllables or so; it is thus a prosodic unit often corresponding to breath groups (Lieberman, 1967), syntagma (Kozhevnikov & Chistovich, 1965), tone groups (Boomer & Laver, 1968; Halliday, 1963), or phonemic clauses (Boomer, 1965; Trager & Smith, 1951). In its usage here, the term "accent" is equivalent to linguistic "stress" except that (formally) accent level is applied relationally to *all* syllables in a pattern. Hence it includes and orders not only all varieties of linguistic stress such as emphatic, primary, sentence, word stress, etc., but also unstressed and reduced vowels. How many of these levels are to be called accented is then an arbitrary choice. This relational approach is thus neutral with respect to the terminological question of "degrees of stress" to be allowed in English (cf. Chomsky & Halle, 1968; Halliday, 1963; Vanderslice, 1968, 1970) or in any other language.

Application of the Rhythm Rule to Speech

On the theory presented here, the terminal rule postulates the relations holding between relative accent level and relative timing for each syllable in the rhythmic pattern. Thus, for example, if accent levels are assigned to syllables in a string, then syllable durations are fixed accordingly. Alternatively, given the relative duration of each syllable in the string plus location of primary accent, the remaining accents are determined. More generally, given some minimum amount of information concerning accent levels and duration on syllables in a string, one could state (or predict) the remaining accent levels and/or syllable durations.

To illustrate, consider Figure 8a, which shows a string of three syllables (elements) of equal duration, with the last syllable marked for primary accent. (Notice that duration of the last syllable in any string is irrelevant since the rule applies not to syllable durations but to syllable loci, specifically, their vowel onsets.)

To place accent level on each syllable, the rhythm rules are applied as follows. (*a*) First determine the minimal tree required. Three evenly spaced elements require a two-node-level tree to yield four terminal branches, one of which will be null. (*b*) Apply the accent rule to the tree to obtain the sequence of accent levels 1324. (*c*) Reverse levels one and two by the terminal rule to obtain 2314. (*d*) Map this sequence onto the syllable string such that accent level one corresponds to the syllable marked for primary accent. The result of this maneuver, shown above the notes in Figure 8b, places relative accent level on each of the three syllables in the string. The last accent level generated by the rule is null, that is, the rightmost tree branch is not actualized. The resulting accent pattern for the elements actualized, 231, is recapitulated below the words in Figure 8b. A vertical bar marks the boundary between "measures" as in Figure 7; this bar will always precede primary accent. Since primary accent always falls on the first left branch in the right half of the rhythmic tree, any pattern can be represented as falling within two "measures." Thus the outer two vertical bars can be omitted, as they are in all remaining examples. Again, primary accent is not restricted to final position by the terminal rule but can occur anywhere in the pattern, as is seen later.

Now consider the three examples in Figure 9, in which the monosyllabic names *John* or *Bill* in Figure 8 have been replaced by the bisyllabic names *Henry, Mary,* or *Marie,* which require one interval subdivision. Put another way, one additional differential relative syllable duration must be considered. Hence one additional node level is required, which generates the accent-level sequence 15372648 by the accent rule, converted to 25371648 by the terminal rule. With primary accent located on the accented syllable of the last word, the sequence is then mapped onto each sentence to yield the accent levels as indicated above the notes. More generally, a sequence of X elements requires a tree with $2^N \geqslant X$ branches, N being the minimal number of node levels if elements are equal in duration. If any one element is one-half the duration of any other, $N + 1$ node levels are required. If any one element is one-fourth the duration of any other, $N + 2$ node levels are required, and so forth.[7]

Numerals below the notes in Figure 9 give accent levels for elements actualized, in rank order. Since three syllables here are arbitrarily counted as accented, the parentheses identify "unaccented" syllables. Thus the accent pattern for each of the three sentences is the same, 231. These examples illustrate the general principle, seen most obviously in languages like English, that syllable durations are compensatorily adjusted so that accented syllables will tend to fall at equidistant intervals. By contrast, on a concatenative view of syllable timing there would be no need for the duration of the word *saw* to change in the differing sentence contexts of Figures 9b and 9c.

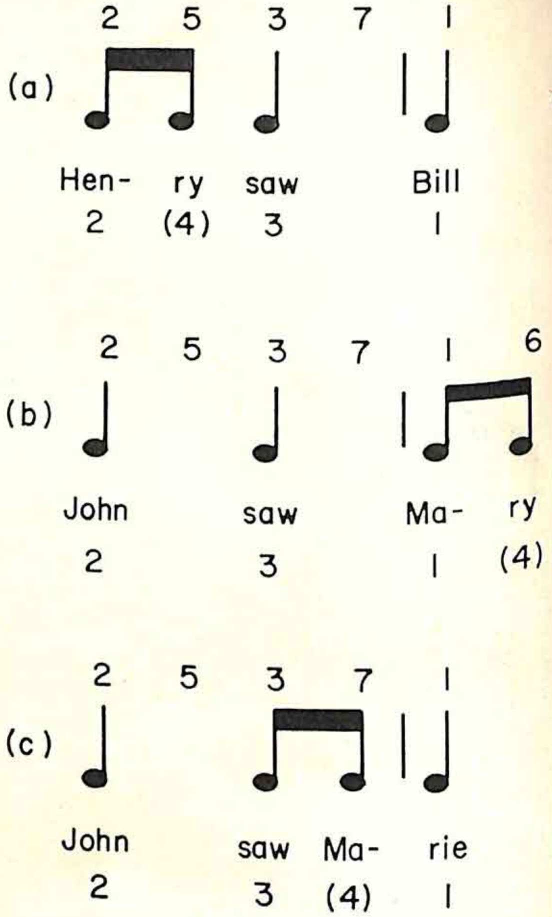


FIG. 9. Applying the terminal rule to three sentences (see text for explanation).

It is important to note that these patterns are somewhat simplified. Accent level and syllable duration are positively correlated (Lehiste, 1970) so that, depending upon how the pattern was pronounced, an accented syllable, for example, *saw* in Figure 9c might be represented as somewhat longer and the following unaccented syllable as correspondingly shorter. At least one additional node level then would be required, in consequence placing lower relative accent on the unaccented syllable (vowel reduction?), but without changing the patterning between the three accented syllables. In principle at least, such details can be represented by the terminal rule whenever required. It is worth noting, incidentally, that each of the sentences in Figure 9 follows the accent pattern assigned by the transformational stress cycle (Chomsky & Halle, 1968).[8] The terminal rule, however, is assumed to be more general, as may be seen next.

---

[7] One additional node level is also required if primary accent falls temporally within the first half of the sequence. In this case, the right half of the tree generated by the above criteria for node levels will not cover the elements in the sequence. One additional node level gives the required result, although now nearly all (sometimes, all) branches on the left half of the tree are null.

[8] The rhythm rule is easily shown to apply also to the idealized pronunciations of words in isolation, although in most cases some intuition about syllable durations is required since phonology gives only notional time. Thus the relation between syntax and stress in English as expressed by the transformational cycle appears to be a special case of the rhythm rule.

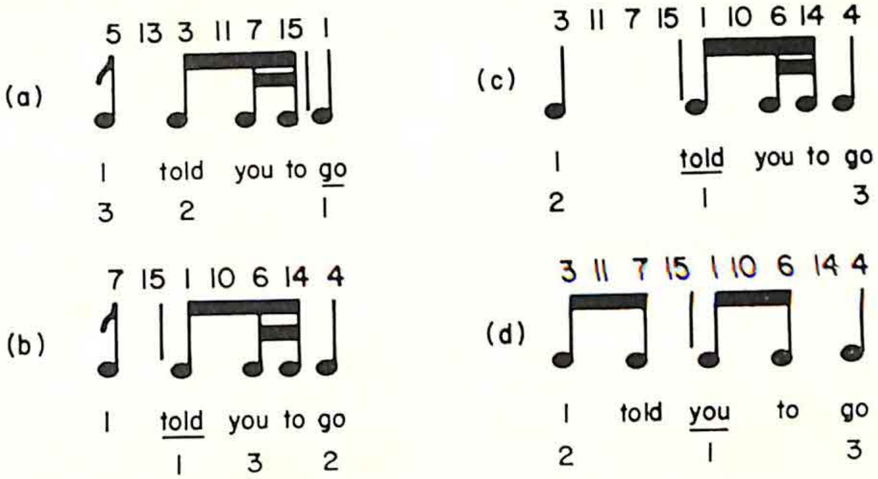


FIG. 10. Four rhythmic versions of the "same" sentence. (The underlined word has been emphasized).

A given sentence, phrase, etc., can be spoken in a variety of ways, being influenced by context, mispronunciation, "foreign accent," regional dialects, the use of emphasis or contrast, and so on. The effects of some of these factors on accent have been discussed by others (e.g., Allen, 1968; Bierwisch, 1968; Bolinger, 1965; Bresnan, 1971; Newman, 1946; Vanderslice, 1968). Consideration of such factors points to another general principle of speech rhythms illustrated here, which is that whenever these factors affect accent placement in an utterance, the temporal pattern of the utterance is affected also. In Figure 8b, emphatic accent on *saw* would give the accent pattern 213; on *John*, 132. To illustrate further, consider the effects of change of emphatic accent in Figure 10. Again the numerals above the notes represent accent level on each syllable by the terminal rule. Numerals below assign accent level in rank order to the three highest values. Although the time values (durations) of the syllables are the same in Figures 10a and 10b, the accent levels, hence rhythmic patterns, differ. Time values are different though accent levels are the same in Figures 10c and 10d. Primary accent is located on the same word in Figures 10b and 10c but the (somewhat unnatural) longer (earlier?) initial syllable in Figure 10c receives greater accent than in Figure 10b, thus changing the whole pattern.[9] Other variations are possible, but in every case a change of accent level or location requires a reorganization of the whole pattern. This principle is assumed to apply to mispronunciation, foreign accents, dialects, and all other factors affecting accent assignment in an utterance.

Notice, finally, that the principle of equal-interval accent holds in all of the examples above except Figure 10a. Given the temporal pattern pronounced exactly as indicated in Figure 10a, the rule requires the accent pattern indicated. If, however, the syllable *I* were shorter than indicated (which is not an unlikely prounciation), the additional required node level would place lower relative accent on *I*, so that the third highest accent would then shift to *you*, giving equal-interval accent pattern 231. On the other hand, given the pronunciation as is, deciding *arbitrarily* in favor of four accent levels rather than three preserves the equal-interval accent principle since then the pattern is 3241. This is a terminological point; the main point is that the tree gives the temporal location of accent however defined (and of "vowel reduction" as well) as it is distributed in natural running speech.

In summary, the rhythmic patterns of speech are wholistic units that can be characterized in terms of the interrelated dimensions of relative accent level and relative timing. These patterns are idealizations; their point-by-point physical correlates will depend on many factors in any actual situation. Notions like these bear strongly on real-time theories of speech production and perception and are considered later in these terms. However, it is appropriate for the moment to consider first some implications for phonetic description.

---

[9] Put another way, placing stronger accent on the initial syllable of Figure 10c requires greater relative duration in this rhythmic context.

Phonetic Considerations

A survey of phonetic research shows that interest in the prosodic characteristics and temporal organization of speech has grown rapidly over the past several years. Earlier the emphasis was predominantly on the segmental characteristics of speech, and questions concerning timing and prosodic features arose more rarely. Three of these questions are considered here. They are relevant because they concern speech timing, but also because they more or less directly illustrate the consequences of adopting one point of view or another about the relations holding between the elements in a string of speech sounds. A given point of view on this problem, whether explicit or not, implicates any number of assumptions about the "units" in speech, research questions to be asked, observations regarded as relevant, perceptual expectations, and so on. Thus the relative timing point of view, which assumes the rhythmic pattern as the unit of temporal organization, entails certain expectations concerning the phonetic manifestations of accent and timing, which differ from those that might be expected on a concatenative, or a neutral, point of view. These differences can be seen in each of the problems discussed below.

Acoustic correlates of accent. A conventional view concerning prosodic and segmental levels of speech, made explicit by Fry (1968), has been to regard segment strings (hence syllable strings) as in some sense the "vehicle" for prosodic phenomena. This view implies, at least by omission, a concatenative view of the temporal relationships between the elements in a string. From this it would seem to follow that when perceived accent occurs in an utterance, its physical correlates are to be located "on" the accented syllables themselves. Accordingly, much of the research on the question of the acoustic correlates of accent typically has been concerned with comparisons based on syllable pairs (e.g., *OBject, obJECT*) without running context, or at least without consideration of the effects of running context. Work in a variety of languages has shown that higher fundamental frequency, greater intensity, and greater duration are generally associated with accented, as opposed to unaccented, syllables, but that there are no simple correspondences between any one or a combination of these acoustic characteristics of a syllable and relative accent (Fry, 1968; Ladefoged, 1967a; Lehiste, 1970; see the latter for an extensive review of the research on prosodic phenomena).

By contrast, the key assumption of the present view is that the rhythmic organization of speech requires accented syllables to fall at roughly equal intervals in time. As shown earlier, this constraint entails that syllable durations, on accented and unaccented syllables alike, will vary according to utterance-specific context. But, in the present theory, utterance-specific context also affects the intensity and frequency characteristics of the syllables, just as it affects duration, one reason being that frequency and intensity characteristics are probably partly determined by the number and rate of articulatory transitions required between syllables in the course of producing a given utterance. Thus the relative values of these characteristics on accented syllables will vary with local context and should not be expected to correlate with perceived relative accent. Instead, the invariance holding between accented syllables is relative timing. Some additional minimal aspects of a rhythmic pattern are assumed to be acoustically marked in the syllable string; these often will include the end of the pattern, which is marked by pause or characteristic fundamental frequency contour, and primary accent (Cushing, 1969; Fonagy, 1966). This information is sufficient to "anchor" the pattern so that relative nonprimary accent on syllables can be "heard" in terms of the rhythm rule, that is, in terms of their location in the temporal pattern. Relative accent is thus a perceptual consequence of relative timing.

It seems clear that accurate reports of accent, in this view, would require some sort of perceptual orientation toward the whole extent of the temporal pattern. For instance, such an orientation is necessary in order to transcribe the temporal pattern of a speech or musical phrase into rhythmic notation once it is heard. But the conviction is strong that the acoustic manifestations of accent are "on" the accented syllables themselves, as attested to by the research devoted to the problem, and indeed it is deceptively easy to attempt comparative judgments based on pitch, loudness, etc. As Lieberman (1965) has shown, however, these judgments will often be wrong, even when they are made by sophisticated linguists. Judgments like these apparently are influenced by a variety of linguistic expectations, with conflicting expectations leading to conflicting judgments. In short, in the present view, relative timing is the proper basis for the perception of relative accent, but it cannot easily be separated from acoustic and other dimensions on which judgments of accent can be, and are, made. A statistical correlation between perceived relative accent and relative timing might emerge over a variety of listeners, listener expectations, and utterances, but would not be highly probable in any particular instance.

Some recent experiments, however, have separated timing from other aspects of running speech, and they show quite clearly that accent and timing are related. These experiments used connected speech materials in which temporal relationships had been artificially distorted. In one approach, judgments of accent on a syllable were shown to depend upon location of the syllable within the rhythmic context of an utterance when the effects of frequency, intensity, and duration of the "accented" syllable itself were ruled out. Listeners heard two versions of sentences and judged relative accent on the last two (monosyllabic) words. Tape splicing prior to the words placed one or the other of the words on the rhythmic "beat." Judged words were thus acoustically equivalent in the two versions. The word on the beat was judged more often as the accented word of the pair (Martin, 1970a).

In a very recent report, using extremely precise electronic splicing techniques, Huggins (1972a) artificially shortened or lengthened a single segment in naturally spoken sentences, for example, the first /m/ in *The hostel for paupers has two moody managers.* Listener reports indicated that as the segment shifted between shortened and lengthened, primary accent shifted between *two* and *managers.* These experiments clearly separate timing from other aspects of the stimulus and show that perceived accent does depend in part at least upon temporal context, and not solely upon the characteristics of the accented syllables themselves.

In another recent paper, Huggins (1972b) artificially varied adjacent segment durations in normally spoken sentences to study compensation effects between them and found that the temporal distortions were more likely to go unnoticed when the original timing relationships between the accented syllables in the sentences remained intact. Again, this experiment demonstrates the importance of relative timing in the perception of accent. Finally, it should be mentioned here that accent is seen as having other correlates than the acoustic characteristics already mentioned, for example, certain measures of physiological effort (Ladefoged, 1967b). Fonagy (1966) regards accent as based upon cues to effort that can be perceived by a listener, even without knowledge of the language. The present view that rhythm, hence accent, is based upon the constraints on natural movement appears to be somewhat compatible with these views.

"Stress-timed" versus "syllable-timed" languages. In the present view, natural constraints on movements entail that running speech must be rhythmic as defined here, so that accented syllables are generally separated from each other and will fall at roughly regular intervals. Obviously, this view also implies that equal-interval accent is a language universal. But the principle of equal-interval accent appears to be perceptually more convincing in the case of some languages than in the case of others, and in fact some languages have been classified as "stress timed" (equal-interval stress) in the former case as opposed to "syllable timed" (equal-length syllables) in the latter (Pike, 1945). Not many phoneticians seem to have accepted this classification, but it is nevertheless of

some interest to consider what might have motivated the choice of this particular dichotomy. It appears that the rhythm rule directly implies the perceptual phenomena upon which the classification is based. To illustrate, assume for the moment, as does the present theory, that all languages are stress timed, in the sense in which the term is used. When word stress in a language is fixed such that widely varying numbers of affixes and function words must be accommodated between stressed syllables, as is the case in English, so that the result is often the "crushing together" (Pike, 1945) of a large number of unstressed syllables, then the perceptual prominence of stressed syllables should be enhanced. This result is implied by the rhythm rule since subdivision of any interval in the pattern adds a node level to the tree, thus increasing the relative difference between the highest and lowest values of accent level placed on each syllable in the pattern. But when word stress in a language is relatively free, or is fixed on syllables at or near a word boundary (either one or the other boundary but not both, in a given language), then the number of syllables between stressed syllables will be fewer, on the average, than in a language like English, hence requiring less interval subdivision and resulting in the impression of evenly spaced syllables and less prominent stress at the same time. Thus English is called stress timed, while French and Spanish, which have late word stress (Delattre, 1965; Harris, 1969; Schane, 1968), have been called syllable timed (Abercrombie, 1967; Harris, 1969).[10] There are, of course, numerous other accent-related linguistic characteristics peculiar to these languages or to any others; but in the present theory these characteristics will sometimes enhance, sometimes weaken, the perceptual prominence of the underlying rhythmic structure that is always present. In short, stress timing is simply more obvious in some languages than others.

Some unpublished data from a study by the author support this view. Very briefly, the purpose (in part) was to show that some unlikely language candidates (viz., Spanish and French, as well as Italian) could be heard as stress timed under suitable conditions.

Tape-recorded excerpts of a monologue provided the stimulus materials in each language. In an indirect approach (*a*) musically trained listeners provided a rhythmic notation, hence temporal pattern, for each utterance. Then (*b*) fluent (mostly native) speakers of each language listened and reported accent patterns (three levels of relative accent on three specific syllables). Application of the terminal rule to the temporal patterns reported by the musicians then predicted accent location and level reported independently by the fluent speakers. The results were that when fluent speakers agreed with each other on which three syllables were accented, the syllables tended to be equidistant, as the theory predicts. Also noteworthy was that the patterns of relative accent on the three syllables followed the terminal rule significantly more often would be expected by chance.

Perceptual and physical time. The claim that English is stress timed has led to several attempts at physical measurement of interaccent intervals (see Allen, 1968a, 1968b; Bolinger, 1965), but these measurements have shown substantial variation between the intervals. In the present view, there are two reasons why a lack of correspondence between perceptual and physical measures of these intervals might be expected. The first is that tempos can change within a rhythmic pattern. When they do, these changes can usually be apprehended yet do not ordinarily affect the perceptual integrity of the pattern, that is, "beats" are still heard as beats; nor do they affect the listener's ability to "follow" or to anticipate later elements in time, since the rhythm is definite though not regular, in Lashley's (1951) words. These tempo changes would be revealed by systematic and cumulative changes in physical syllable durations, either increasing or decreasing, depending upon whether the change was positive or negative. But to detect these systematic changes and separate them from measurement error in any given utterance, it is necessary first to postulate a specific rhythmic pattern, that is, the pattern of syllable durations that would obtain in an even tempo, and then to compare departures from these values in the observed durations by curve-fitting or similar routine. For instance, relative time intervals from an utterance heard and transcribed in rhythmic notation might be plotted against cumulative measured vowel onsets. A straight line would indicate even tempo, a curved line would indicate an *accelerando* or *ritard.* Clearly, alternatives such as calculating "average" or "intrinsic" syllable durations without regard to specific pattern context would be of little value in this regard.

---

[10] Note also that languages in which accent is located relative to word boundaries are unlikely to have a transformation cycle relating accent to syntax, since there is little of syntactic significance in a word boundary.

A second reason for expecting a lack of correspondence between physical and perceptual measures of the interval between accents involves interesting perceptual problems that have been little explored. One of these concerns the correlation between perceptual and physical measures of the dimensions of a rhythmic or any other auditory temporal pattern when the dimensions are considered independently as opposed to jointly. There is some evidence that the time dimension interacts with the accent-level dimension so that physical and perceptual measures of either dimension taken alone will not correspond one to one. Woodrow (1909), using nonspeech stimuli, showed that variations in physical relative intensity level on elements affect perceived relative timing. On the other hand, variations in physical relative timing affect perceived accent level (Martin, 1970a), as mentioned earlier.

Rhythmic Structure in Speech Production and Perception

Several Hypotheses about Speaking and Listening in Real Time

Speech utterances in the present view are temporally patterned, and many of their elements are temporally predictable. These characteristics of the auditory stimulus constrain the class of mechanisms that might be postulated for their efficient processing (Garner, 1970b). In this section, main interest centers on how linguistic information might be encoded into and decoded out of the speech signal at the rhythmic (prosodic) level, but some implications for the segmental level are considered also. As will be seen, there is reason to believe that the importance of prosodic as opposed to segmental factors has been underestimated in theories of the production and perception of everyday connected speech.

Hierarchical organization of the speech production program. Any evidence for rhythmic patterning in speech, that is, that the syllables in a string are timed relative to the whole pattern rather than concatenated, is also evidence for preset timing and hence leans toward the view favoring central rather than peripheral timing mechanisms during production.[11] A more specific hypothesis which follows is that since the accented elements dominate the temporal organization of an utterance, they must in some sense be planned first. Intervening, lower-level syllables then are planned subsequently in hierarchical fashion, by (metaphorically) reading the rhythm tree, level by level, from the top down. Planning here means at least selecting the time at which syllables will be actualized, although this decision must facilitate any number of other contingent decisions during production. In this view one might think of accented syllables as the main targets in the organization of the articulatory program. The hypothesis then predicts that decisions concerning accented syllables will precede those concerning unaccented syllables, hence that decisions concerning "content" words will precede those for "function" words, and so forth. Observed utterances do not often illuminate the mechanisms generating them; but errors, for example, slips of the tongue, sometimes do. Fromkin's (1971) interesting work on slips of the tongue revealed that whole words in an utterance could be transposed without affecting the overall accent pattern of the utterance, for example,

$$\underset{2}{\underline{\text{a laboratory}}}\ \underline{\text{in}}\ \underline{\text{our}}\ \underset{3}{\underline{\text{own}}}\ \underset{1}{\underline{\text{computer.}}}$$

---

[11] Invertebrate physiologists appear virtually to take for granted that rhythmic behavior is predominantly under central control (see many papers in Wiersma [1967]).

Her interpretation of this result was that the syntax and the stress contour (accent pattern) of the utterance was generated prior to word selection. There are, however, only a few accent patterns, of which 231, the pattern in all but one of Fromkin's three-accent-level examples, is the most common. On the present theory, sufficient decisions about content words are made so that they can be ordered (or misordered), and the syllables marked for accent mapped onto the higher branches of the rhythmic tree. Function words, etc., lower down on the tree are organized subsequently. Thus reversals occur between elements at the same level, and reversals between a content and function word would not be expected (unless the function word was of the often-accented dangling participle type, e.g., *He called her UP*). Other data provide a bit more detailed support for the top-down generation hypothesis. Boomer and Laver's (1968) analysis of slips of the tongue in tape-recorded speech showed that reversals involved a pair of accented syllables most often, a pair of unaccented syllables less often, but rarely an accented and unaccented syllable. Alternatives to the top-down hierarchical generation hypothesis cannot be ruled out completely, but if they are to be viable they should be expected to take into account the priorities that accented syllables appear to have in production.

Finally, Goodglass, Fodor, and Schulhoff (1967) reported that certain aphasics had difficulty imitating unaccented words in three-word phrases when the unaccented words were phrase initial (*do BIRDS FLY*) but not when they were phrase medial (*DOGS can BARK*). They concluded that prosodic rather than grammatical characteristics of function words determined whether they were lost or retained in ungrammatical speech and further that these aphasics depended upon accented features of an intended utterance for initiating and maintaining speech. These data and conclusions, which imply "momentum" in speaking, are consistent with the present view that accented syllables are the articulatory targets of ballistic movements, and that intervening unaccented syllables are produced by "secondary" articulatory gestures "en route" to the target syllables.

Syntax and morphology in rhythmic patterns. Probably all languages provide for the alternation of accented and unaccented syllables in their morphological and syntactic structure, as in English in which "function" words and affixes intervene between the ordinarily accented words and syllables. To the extent that syntactic and other information is encoded at the rhythmic level, such information becomes available for perception independently of analysis at the segmental level. While it is too early to say whether highly abstract syntactic relations will be found to correlate with the prosodic contours of a rhythmic pattern, there are many perceptual decisions about syntax, morphology, and meaning that can be made on the basis of syllabic contour alone. For instance, the names *Mary* and *Marie* in Figures 9b and 9c could be distinguished from each other in the absence of clear segmental detail. Recognition that *John was hit by Bill* is a passive, not an active, sentence requires no decisions concerning the phonetic details in the function words. The syllabic contour of *John flew to WxYz*, in which *W* and *Y* are accented syllables, and *x* and *z* are unaccented syllables, rules out Boston and New York as possible destinations, as well as Minneapolis, assuming the latter is correctly pronounced, etc.

While these are simple examples, there is a good deal of evidence that rhythmic patterning carries a heavy information load in ordinary connected speech. This is particularly evident in cases where speech remains intelligible when segmental information has been distorted or destroyed. Most of the following studies did not refer to rhythmic patterning, and none gave explicit patterns, but the nature of the distortions in each of them was such as to leave the temporal, hence rhythmic, patterning of the signal intact. One obvious example of this from everyday experience is whispered speech, a distortion affecting mainly the frequency and amplitude but not time dimensions of the signal. The fact that whisperers apparently compensate by exaggerating stress (Fonagy, 1968) shows the effectiveness of rhythmic factors in communication.

Cherry and Wiley (1967) and Holloway (1970) gated "weakly-voiced segments"

(Holloway, 1970) out of running speech, replaced these destroyed segments with white noise, and found that the speech again became intelligible. Both papers proposed rhythmic continuity as one aspect of the signal left unaffected although, clearly, segmental information had been removed. While the result is hardly surprising since the missing segments are just those that would be masked in noisy environments, it does show the contribution of prosodic features to intelligibility.

Of the various distortions that nevertheless do not destroy speech intelligibility (Licklider & Miller, 1951), one of the most severe is infinite peak clipping. Two musicians in the author's laboratory transcribed in rhythmic notation a passage of running speech distorted in this fashion.[12] Their rhythmic patterns agreed strongly with each other and with a subsequent transcription of the corresponding normal speech made later by one of them. With unlimited time and repeated listening, they made few errors of word identification, but when they did the errors were usually rhythmically identical to the original words. Lindner (cited in Bondarko, Zagorujko, Kozhevnikov, Molchanov, & Chistovich, 1970) has reported that spectrally distorted speech also retains its rhythmic pattern.

Blesser (1969) transformed the speech signal by rotating spectral energy around a central frequency of 1,600 Hertz [Hz.] so that, for example, energy at 2,000 Hz. became energy at 1,200 Hz. While this transformation distorts segmental information, temporal patterning and fundamental pitch contour remain unchanged. Pairs of subjects in isolated rooms learned to converse with each other rather successfully, hearing only the transformed speech of the other subject. It is tempting to suppose that their communicative success depended upon learning the transformation, but this was not Blesser's conclusion. Among his findings were that intelligibility scores on isolated syllables, words, and sentences did not correlate with communicative success, and that misheard sentences were often syntactically (and rhythmically) nearly identical to the input sentences but otherwise completely incorrect (e.g., *Hoist the load to your left shoulder* was reported as *Turn the page to the next lesson*). Among the conclusions Blesser drew from his study was that syntax was encoded into prosodic features, but more generally that intensive analysis of phonetic details would not solve the problems of speech segmentation and recognition. Studies like these appear to show that there is extensive information available in speech at the rhythmic level that can be used by the listener. They do not show, of course, that he ordinarily does or must use the information under ordinary listening conditions, but it does not seem wise to conclude that the information is used only as a last resort. Future research will determine the levels of complexity and detail of the syntactic and other information that are or can be encoded into the rhythmic contour of the speech signal. Conceivably some languages may make more informative use of rhythmic constraints on the sound patterning of speech than others.

Rhythmic deployment of phonetic detail. The hypothesis that accented syllables are the main targets of ballistic movements during the course of production suggests one possible consequence that also has implications for perception, namely, that accented syllables are articulated at the expense of lower-order intervening syllables in fast speech, or perhaps even normally. The more syllables intervening (subdividing) between accented syllables then, the greater the extent that phonetic detail on these syllables will be blurred. In music, a situation analogous in some respects exists when the performer, encountering a passage too fast to articulate, takes pains to get the notes on the beat and simple multiples thereof, slurring over the unaccented intervening notes, or even omitting them completely. The melodic outline usually is preserved nonetheless, and the slurred or omitted notes may even escape notice. In the case of speech, it is known that formants of weakly stressed vowels depart from "target" values, approaching a neutral value (schwa) during fast speech (Lehiste, 1970; Lindblom, 1963). These

---

[12] Gratitude is expressed to John Boehm, who made the recording.

variants on pronunciation contribute to the well-known enormous variability of the acoustic correlates of a given phoneme (MacNeilage, 1970) in ordinary connected speech. But schwa is uninformative concerning the phoneme "intended" by the speaker, and presumably many other vowel pronuciations are also, the more so that they approach schwa. These variants on the acoustic correlates of phonemes complicate the perceptual problem for any mechanism that approaches the signal as though every interval in the signal potentially contains useful segmental information, and hence that every interval must be taken account of to determine whether useful information is present or not. There is no evidence, however, that speech perception proceeds this way. The alternative view proposed here is that the perceptual analysis is optimized by prior information as to the more informative stretches of the signal, for example, on-target vowels plus preceding consonants, which appear to be programmed with them (Tatham, 1970). On the present rhythmic hypothesis these informative stretches, namely, accented syllables, are systematically related to the temporal dimension of the signal. Analysis of the utterance would consist in determining the rhythmic pattern of the syllable string by means of a wholistic (Neisser, 1967) "preliminary analysis" (Halle & Stevens, 1964), then analyzing for segmental detail by reading down the rhythmic tree, so that the more veridical stretches of the signal are analyzed first, with decisions about the lower-level details made conditional upon the results of the higher-level analysis. Such a routine is made possible by the fact that the rhythmic patterning can be extracted from the signal independently of a segmental analysis.

This hypothesis is made plausible by evidence that speech is rhythmically patterned, but it would be strengthened by evidence that accented syllables (and segments therein) are easier to recognize than unaccented syllables (and segments therein) when these two kinds of syllables are removed from linguistic context. In a study using such an approach Lieberman (1963) placed experimental words in two sentences each such that the word was informative (presumably accented) in one sentence or redundant (presumably unaccented) in the other, and gated the words out of the running speech context. Finding that the excised informative words were recognized more easily than the redundant words in listening tests, he concluded that they were more clearly articulated than the redundant words.

Another way of excluding linguistic context from elements to be recognized is to use listeners ignorant of the language. In the language study mentioned earlier, the pair of musicians who transcribed the Italian speech in rhythmic notation were also required to first transcribe the sounds even though they did not know Italian. Analysis of the transcriptions showed more agreement on accented than on unaccented vowels, as they were recorded in English alphabetic notation. These listeners of course were locating the sounds in coarse categories within the perceptual space of their own language. Presumably, phoneticians would often report schwa in the unaccented syllables, hence agreeing more often on these syllables, whereas the musicians, who knew nothing of phonetics, drew their guesses from a pool of English vowels. The relevant point here is that differential agreement appears to be based upon information tied to the rhythmic patterning of the utterances. These results and Lieberman's (1963) seem to be consistent with the hypothesis that the deployment of usable segmental information in the speech signal is temporally systematic. If the hypothesis is correct, future research may determine if and how the temporal information is ordinarily used in perceiving the speech.

Dynamic coupling of speaking and listening in real time. The hypothesis of the preceding section, which suggests a two-stage analysis with the segmental stage depending upon the preceding prosodic stage, bears some resemblances to the analysis-by-synthesis model of speech perception (Halle & Stevens, 1964), as implied above. The analysis-by-synthesis model postulates a preliminary analysis of the input signal prior to an (active) analysis-by-synthesis matching routine. The input signal which is the object of the preliminary analysis has not usually been specified, but could be the rhythmic pattern as implied above. Of particular relevance here is that whatever the size of the input signal, or "decision unit" (Miller, 1962), a sequential stage model appears to imply that the last stage of the analysis, and possibly even the preceding one, could not proceed until the end of the unit had been reached, since some terminal cue would seem to be required as the sign to start analyzing. In running speech this routine thus appears to amount to a continuous, repeating sequence of (*a*) accept input acoustic signal unit, (*b*) apply (first stage) preliminary analysis, (*c*) apply (second stage) matching or other routine, (*a*′) accept next input acoustic signal unit, (*b*′) . . . , etc. Even if Stages *a* and *b* overlapped in time it would still be true that the smaller the decision unit (e.g., syllable) the more rapidly and often the analytic stages would be required to alternate, whereas the larger the unit (e.g., rhythmic pattern), the faster the last stage would be required to operate. An alternative possibility, which would permit a large unit yet more leisurely analysis, is to suggest that Stage *a* above proceeds independently of and prior to Stage *b*, and that Stages *b* and *c* operate sequentially with respect to each other but simultaneously with the *following* Stage *a*′. That is, the two-stage analyses of a preceding pattern are distributed over the time during which a following pattern is simultaneously being accepted into acoustic memory. This notion, however, seems counterintuitive since we appear to understand speech as we hear it.

The rhythmic point of view provides a somewhat different alternative, which may be explicated as follows. As in the hypothesis of the preceding section, a prosodic and segmental analysis is assumed. Inherent in the rhythmic concept is that the perception of early events in a sequence generates expectancies concerning later events, in real time. When the events are sounds produced by continuous movements, the perception of these includes cues as to the movement dynamics involved in their production. Hence it is not simply, or not only, that discrete arrival times of accented syllables are induced from earlier timing relationships but also that the total array of time-varying cues in the continuous flow of speech will project ahead the general outline of the remaining prosodic contour. These cues telegraph not only tempo changes but more generally the whole thrust of the pattern of sounds yet to come. It is on this basis that one might say not that the listener "follows" the speaker but rather that the listener, given initial cues, actively enters into the speaker's tempo. Whereas the analysis-by-synthesis model is a feedback model (Halle & Stevens, 1964), the rhythmic model presented here appears to have also a feed-*forward* aspect (MacKay, 1967), in the sense that it adjusts ongoing processing based upon information about future, as opposed to present or past, states of the signal.

The consequences for running speech perception in this view are that temporal redundancy at the prosodic level permits a prosodic analysis to begin prior to the end of the decision unit. This in turn might even permit an earlier start toward segmental analysis, since the latter depends upon the results of the prosodic analysis. The temporal redundancy also permits accented syllables to be anticipated in real time, so that "attention focusing" on accented syllables becomes a logical possibility, as does "attention cycling," such that previous inputs could be processed during the low-information intervals between accented syllables. Real-time mechanisms like these appear to be of the sort needed to account for the fact that speech processing is so rapid (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).

There is at present very little evidence in the literature that would bear on these considerations, although two studies can be mentioned. These concern the role of accented syllables in perception. Allen (1967) reported that tapping variances were smaller to accented syllables than to unaccented syllables when the subjects were required to tap in time with a single specified syllable in utterances presented repeatedly by means of a tape loop. Since it would seem that the arrival times for both kinds of target syllables were equally predictable in the sense that the subject had heard the utterance many times and knew "where" the syllable was in the utterance, the results may have to do with the ease with which the listener can coordinate his activities with those of the speaker. In the present view, the listener taps more accurately when his tapping targets are the same as the speaker's speaking targets, that is, accented syllables. The report by Huggins (1972b) mentioned earlier, which showed that listeners were insensitive to certain artificial distortions in speech when the original timing relationships between accented syllables were preserved, is also consistent with the view presented here. Further research may determine the extent to which efficient perceptual processing involves actively locking into the temporal flow of the speech signal.

Applications to Automatic Speech Synthesis and Recognition

These research areas are concerned either with natural running speech as inputs or with natural sounding speech as outputs. Here some suggestions are offered as to ways in which rhythmic structure considerations might be relevant.

Speech synthesis by rule. At the present time, speech synthesis programs typically concatenate segments whose durations are determined by dictionary look-up, with adjustments in duration based upon immediate segment context (see, e.g., Flanagan, Coker, Rabiner, Schafer, & Umeda, 1970; Mattingly, 1971; see also Denes, 1970; Mattingly, 1966; Rabiner, 1969; Rabiner, Levitt, & Rosenberg, 1969; Teranishi & Umeda, 1968). In the present theory, the temporal patterning of utterances constructed this way can be expected to deviate from those of natural speech. Additionally, since dictionary look-up durations presumably will always be of sufficient length so as to permit some phonetically differentiating detail, these details will be distributed relatively evenly throughout the utterance so that the resulting utterance contains more phonetic detail, or at least more evenly distributed detail, than the same utterance in natural speech. The requirements of relative timing in natural speech, on the other hand, often will obliterate phonetic detail when unaccented syllables are crowded together in utterance-specific rhythmic contexts. Possibly such deviations from natural speech account in part for the reported trade-off between naturalness and intelligibility in synthesized speech (Mattingly, 1971).

In principle at least, these speculations could be rather directly tested by experiments on synthetic speech. The general approach would be to mark syllables to be accented in the syllable string and then apply the rhythm rule to the string. Relative syllable durations would thus be determined, and hence also the phonetic details affected by duration. Whether such rhythmic constraints result in synthesized speech which is easier to listen to over long periods of time or is of better judged quality is then an empirical question to be settled by listening tests of various kinds, including tests that would be sensitive to long-term fatigue effects, etc., hence to the "cost" of processing distorted inputs.

Automatic speech recognition. Automatically synthesized speech can be less than perfect yet remain intelligible because of the redundancy in speech and the perceiver's ability to compensate for the inadequacies of the signal. It is the latter that allows judgments to be made about an inadequate input to the listener, so that trial-and-error procedures can be used to evaluate progress toward better quality synthesis. Automatic speech recognition, on the other hand, is seen to be particularly difficult because speech recognition by humans depends so little upon information in the acoustic signal and so heavily upon knowledge of the language and of the world not possessed by computers (Pierce, 1969). Nevertheless, the rhythmic hypotheses considered earlier suggest an approach that in principle at least might make it possible to extract more information from the acoustic signal than is now the case, and additionally, bring the perceiver's abilities into the automatic recognition routine. This approach involves both the analysis of the signal and the resulting printout. The first stage of analysis would consist in determining the syllabic pattern, by locating vowels or vowellike segments.

The second stage would consist in extracting the more veridical segmental information by means of a top-down analysis as suggested earlier. The printout would reflect the two-stage nature of this analysis by providing a symbol sequence such that the syllabic pattern is recognizable regardless of the accuracy of identification of any of the segments therein. Identifiable segments (those exceeding some cutoff decision criterion) would be represented by appropriate phonetic symbols. Unidentifiable vowellike segments might be represented by schwa; unidentifiable consonantlike segments might be represented by some neutral symbol. Such an output provides a sequence in which the general syllabic pattern may be made out although many of the details are blurred. But such a sequence, like poor ditto copy or poor handwriting, might be more or less readable, since strategically located and correctly identified consonants might cue whole words. Thus by bringing the perceiver's (reader's) knowledge of the language and of the world to bear on evaluation of the printed output, trial-and-error procedures become available to the speech recognition worker that are quite analogous to those enjoyed by those working in speech synthesis. And, possibly, the computer recognizing speech might require little more knowledge of the language and of the world than that required by the computers assigned to synthesize speech.

FURTHER, FINAL SPECULATIONS CONCERNING OTHER RESEARCH AREAS

Rhythmic structure considerations are potentially relevant in a variety of research areas. In this concluding section, several of these areas are mentioned. In some cases the relevance of rhythmic considerations is fairly clear, but in others it can be only highly speculative.

Immediate Memory

Work in this research area has depended heavily upon the use of digits, letters, and other spoken stimulus inputs. Adams investigated rhythmic factors in immediate recall of digit strings as early as 1915, but there has been relatively little interest in the rhythmic aspects of the speech stimulus in memory experiments until quite recently (e.g., Neisser, 1967).

Most contemporary experiments on the role of temporal factors in immediate memory have tended to focus on the grouping aspect of rhythmic sequences, using stimulus inputs designed to provide grouping by the insertion of pauses into the sequences. The effect of pauses has been seen as that of providing perceptual groupings and of providing the opportunity for coding, rehearsal, etc., during the pause interval (e.g., Bower & Springston, 1970; Laughery & Spector, 1972). The present view would go further than this, to assume that speaking digits and other verbal input strings requires an organized articulatory program just as does normal speaking and hence produces a structured stimulus to which the listener responds in the manner discussed earlier. Many elements of this structure persist (it is assumed) even when the speaker tries to speak without intonation. Factors like these are often irrelevant to the purposes of the memory experiment, but they will nevertheless play a role in ongoing processing of the stimulus string. Sometimes these factors will be particularly important, for example, when group sizes vary such that the subject does not know in advance how many items he will hear in a group (e.g., Bower & Winzenz, 1969; Laughery & Spector, 1972), in which case prosodic cues on group-initial elements could project ahead the general prosodic contour of the group, which includes cues as to the number of items it contains. The contribution of perceptual factors like these to memory experiments has not been systematically assessed, but there is evidence that they exist, as the following experiment shows.

Neisser, Hoenig, and Goldstein (1969) examined the "prefix effect" from a point of view quite similar to that of the present paper. Earlier, Dallett (1964), following Conrad's (1958, 1960) work, had shown that a redundant stimulus prefix spoken at the beginning of a seven-digit string reduced recall of the digits, even though the subject was not required to repeat the redundant digit, as compared to the condition in which the spoken redundant digit did not precede the string. Neisser et al. (1969) showed that the prefix effect disappeared when the redundant digit was uttered by a different speaker than the one speaking the digits to be recalled, so that the prefix could not be a part of the same production program. It appears that the listener actively responded to the whole spoken pattern, which, as the experimenters pointed out, required him to handle an eight-digit string at input and suppress the prefix before speaking the recall digits back. Hence recall of prefixed seven-digit strings yielded about the same percentage recall as eight-digit strings. This experiment seems to show that intonation contour or other aspects of the speech stimulus can affect recall, but it does not make clear which details are important. Probably a systematic analysis will require comparisons between normally spoken speech strings and other strings with elements timed the same way but with articulatory continuity removed by tape splicing or synthetic methods. It may even be possible to demonstrate that natural articulatory grouping can, under certain conditions, reduce or even eliminate the conventional advantage of pauses within input strings to be recalled (Martin, 1969a).

Serial Pattern Learning

Restle (1972) investigated the learning of serial patterns which, in contrast to materials used in the section above, have an inherent structure as described by structural trees (Restle, 1970; Restle & Brown, 1970). Pauses introduced during presentation faciliated pattern learning, as is typically the case in serial learning, but only when the pauses were appropriate for the internal structure of the pattern; that is, they separated subunits. It is a reasonable guess that learning of these meaningful sequences might be further facilitated if the timing of elements in a sequence during input not only marked subunits but also was temporally organized around an accent structure following the accent rule. It should also be noted that Restle's (1970) theory has been shown to be related to aspects of music structure.

Auditory Temporal Pattern Perception

In the present theory, natural auditory temporal patterns, speech or otherwise, are sound sequences with internal structure as described by the rhythm rules and with a wholistic unity. Hearing early elements in a pattern generates expectancies concerning later elements, in particular, the temporal location of accents. For instance, accents on elements one and three and/or five in an eight-element evenly spaced pattern follow the accent rule, whereas accents on elements one and four do not follow the rule and lead to expectancies of triplets, which are appropriate for nine-element patterns but inappropriate for eight-element patterns. Expectancies that are confirmed should make any kind of perceptual processing easier, whereas expectancies that are violated should make processing harder.

Most of the extensive work by Garner and his associates (Garner, 1970a; Garner & Gottwald, 1968; Preusser, Garner, & Gottwald, 1970; Royer & Garner, 1966, 1970) on auditory temporal patterns does not bear directly on this hypothesis since their methods were directed toward determining preferred organizations. But one of their findings was suggestive for the present theory. In these experiments, after hearing several continuous repetitions of a pattern, the subjects, whose task was always to report the pattern correctly, generally preferred to describe or track the patterns by placing runs either at the beginning or end of the pattern, regardless of the starting point of the pattern as it was presented to them. These were their preferred organizations of the pattern. Of interest here, however, is the case where there were no runs of substantial length, so that no run preference could apply.

In Royer and Garner's (1966) experiment, there were no runs longer than two elements in patterns R, S, and T. Their data show that for each of these patterns, response delays were less, on the average, when the starting point presented to the subjects followed the accent rule (i.e., placed accents on elements one and three) than when they did not follow the rule (i.e., placed accents on elements one and four). It appeared that when the patterns did not have the characteristics that permitted preferred organizations (two of these three patterns were the most difficult in the experiment), then the effects of starting point made a difference in the time required for the subject to confidently begin tracking the pattern. This interpretation is of course highly speculative, and there may be other explanations, but it is supported by evidence from an experiment designed explicitly to test this interpretation. Sturges and Martin (1971) presented patterns continuously but only twice; the second pattern was either exactly the same (repeated) or changed slightly. The tasks were to judge same or different and then write the pattern(s). Repeating patterns that followed the accent rule were more easily recognized as repeating than patterns that did not follow this rule. Similar results held for written recall.

Language Acquisition

Prosodic features have been a traditional concern in studies of child language. For instance, children appear to learn intonation patterns before they learn segmental features (Weir, 1966). Surely the rhythmic characteristics of speech will make a difference in the way sentences are heard by children. Take a sentence like *The girl the boy saw went home.* This sentence sounds equally natural if *boy* is more strongly accented than *saw* or vice versa; yet the two pronunciations could lead the child to give different imitations or paraphrases, which in turn would lead the developmental psycholinguist to different conclusions concerning the comprehension of embedded sentences. More generally, it would not be surprising if a great deal of child language data that seem refractory on syntactic grounds would make sense if considered in terms of real-time phonetic considerations.

Meaning

Rhythmic patterning applies to sounds, but is there anything to be said about semantics? Kozhevnikov and Chistovich (1965) have pointed out that the syntagma, their articulatory unit, has been defined also as a unit of meaning. McNeill (1971), who, incidentally, also cited Lashley's (1951) paper on serial order, has advanced the view that speech units contained within certain intonation contours are underlying structure sentences. Perhaps future work will show that ideas and meanings are efficiently encoded into natural rhythms. There is an old tune, from the 1930s or so, called "It Don't MEAN a Thing if it Ain't Got That SWING." It refers, of course, to the rhythms of music, but there may be a sense in which it is true for the rhythms of language also.

REFERENCES

Abercrombie, D. *Elements of general phonetics.* Chicago: Aldine, 1967.

Adams, H. F. A note on the effect of rhythm on memory. *Psychological Review,* 1915, **22**, 289–298.

Adams, J. A. Response feedback and learning. *Psychological Bulletin,* 1968, **70**, 486–504.

Allen, G. D. Two behavioral experiments on the location of the syllable beat in spoken American English. *Studies in Language and Language Behavior,* 1967, **4**, 2–179.

Allen, G. D. The place of rhythm in a theory of language. (Working Papers in Phonetics, No. 11). Los Angeles: University of California, 1968. (a)

Allen, G. D. Towards a description of stress-timing in spoken English. In E. Zale (Ed.), *Proceedings of the Conference on Language and Language behavior.* Ann Arbor: University of Michigan, 1968. (b)

Bierwisch, M. Two critical problems in accent rules. *Journal of Linguistics,* 1968, **4**, 173–176.

Blesser, B. A. Perception of spectrally rotated speech. Unpublished doctoral thesis, Massachusetts Institute of Technology, Department of Electrical Engineering, 1969.

Bolinger, D. L. Pitch accent and sentence rhythm. In I. Abe & T. Kanekiyo (Eds.), *Forms of English.* Cambridge: Harvard Press, 1965.

Bondarko, L. V., Zagorujko, N. G., Kozhevnikov, V. A., Molchanov, A. P., & Chistovich, L. A. A model of speech perception by humans. (Trans. by Ilse Lehiste) (Working Papers in Linguistics, No. 6) Columbus: Ohio State University, 1970.

Boomer, D. S. Hesitation and grammatical encoding. *Language and Speech,* 1965, **8**, 148–158.

Boomer, D. S., & Laver, J. D. M. Slips of the tongue. *The British Journal of Disorders of Communication,* 1968, **3**, 2–12.

Bower, G. H., & Springston, F. Pauses as recoding points in letter series. *Journal of Experimental Psychology,* 1970, **83**, 421–430.

Bower, G. H., & Winzenz, D. Group structure, coding, and memory for digit series. *Journal of Experimental Psychology,* 1969, **80** (2, Pt. 2).

Bresnan, J. W. Sentence stress and syntactic transformations. *Language,* 1971, **47**, 257–281.

Cherry, C., & Wiley, R. Speech communication in very noisy environments. *Nature,* 1967, **214**, 1164.

Chomsky, N. *Syntactic structures.* The Hague: Mouton, 1957.

Chomsky, N. A review of B. F. Skinner's *Verbal behavior. Language,* 1959, **35**, 26–58.

Chomsky, N., & Halle, M. *The sound pattern of English.* New York: Harper, 1968.

Conrad, R. Accuracy of using keyset and telephone dial, and the effects of a prefix digit. *Journal of Applied Psychology,* 1958, **42**, 285–288.

Conrad, R. Very brief delay of immediate recall. *Quarterly Journal of Experimental Psychology,* 1960, **12**, 45–47.

Cooper, G. W., & Meyer, L. B. *The rhythmic structure of music.* Chicago: University of Chicago, 1960.

Cushing, S. English as a tone language: The acoustics of primary stress. (Quarterly Progress Rep. No. 92) Cambridge, Mass.: M.I.T. Research Laboratory for Electronics, 1969.

Dallett, K. M. Effects of a supplementary prefix on recall. *Journal of Experimental Psychology,* 1964, **67**, 296–298.

Delattre, P. *Comparing the phonetic features of English, French, German and Spanish.* Heidelberg: Verlag, 1965.

Denes, P. B. Some experiments with computerized speech. *Behavioral Research Methods and Instrumentation,* 1970, **2**, 1–5.

Flanagan, J. L., Coker, C. H., Rabiner, L. R., Shafer, R. W., & Umeda, N. Synthetic voices for computers. *IEEE Spectrum,* 1970, **7**, 22–45.

Fonagy, I. Electro-physiological and acoustic correlates of stress and stress perception. *Journal of Speech and Hearing Research,* 1966, **9**, 213–244.

Fonagy, I. Accent and intonation in whispered speech. *Phonetics,* 1968, **20**, 177–192.

Fromkin, V. A. The non-anomalous nature of anomalous utterances. *Language,* 1971, **47**, 27–52.

Fry, D. Prosodic phenomena. In B. Malmberg (Ed.), *Manual of phonetics.* Amsterdam: North-Holland, 1968.

Garner, W. R. Good patterns have few alternatives. *American Scientist,* 1970, **58**, 34–42. (a)

Garner, W. R. The stimulus in information processing. *American Psychologist,* 1970, **25**, 350–358. (b)

Garner, W. R., & Gottwald, R. L. The perception and learning of temporal patterns. *Quarterly Journal of Experimental Psychology,* 1968, **20**, 97–109.

Goodglass, H., Fodor, I. G., & Schulhoff, C. Prosodic factors in grammar—Evidence from aphasia. *Journal of Speech and Hearing Research,* 1967, **10**, 5–20.

Greenwald, A. G. Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review,* 1970, **77**, 73–99.

Halle, M., & Stevens, K. S. Speech recognition: A model and a program for research. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language.* Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Halliday, M. A. K. The tones of English. *Archivum Linguisticum,* 1963, **15**, 1–28.

Halwes, T., & Jenkins, J. J. Problem of serial order in behavior is not resolved by context-sensitive associative memory models. *Psychological Review,* 1971, **78**, 122–129.

Harris, J. W. *Spanish phonology.* Cambridge: M.I.T. Press, 1969.

Holloway, C. M. Passing the strongly voiced components in noisy speech. *Nature,* 1970, **226**, 178–179.

Huggins, A. W. F. Just noticeable differences for segment duration in natural speech. *Journal of the Acoustical Society of America,* 1972, **51**, 1270–1278. (a)

Huggins, A. W. F. On the perception of temporal phenomena in speech. *Journal of the Acoustical Society of America,* 1972, **51**, 1279–1290. (b)

Keele, S. W. Movement control in skilled motor performance. *Psychological Bulletin,* 1968, **70**, 387–403.

Kozhevnikov, V. A., & Chistovich, L. A. *Speech: Articulation and perception.* Washington, D. C.: Joint Publications Research Service, 1965.

Ladefoged, P. *Three areas of experimental phonetics.* London: Oxford University Press, 1967. (a)

Ladefoged, P. *Linguistic phonetics.* (Working Papers in Phonetics, **6**, Phonetics Laboratory). Los Angeles: University of California, 1967. (b)

Lanier, S. *The science of English verse.* New York: Scribner, 1880.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley, 1951.

Laughery, K. R., & Spector, A. The roles of recoding and rhythm in memory organization. *Journal of Experimental Psychology,* 1972, **94**, 41–48.

Lehiste, I. *Suprasegmentals.* Cambridge: M.I.T. Press, 1970.

Lenneberg, E. H. *Biological foundations of language.* New York: Wiley, 1967.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review,* 1967, **74**, 431–461.

Licklider, J. C. R., & Miller, G. A. The perception of speech. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

Lieberman, P. Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech,* 1963, **6**, 172–179.

Lieberman, P. On the acoustic basis of the perception of intonation by linguists. *Word*, 1965, **21**, 40–54.

Lieberman, P. *Intonation, perception, and language.* Cambridge: M.I.T. Press, 1967.

Lindblom, B. E. F. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 1963, **35**, 1773–1781.

MacKay, D. M. Ways of looking at perception. In W. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form.* Cambridge: M.I.T. Press, 1967.

MacNeilage, P. F. Motor control of serial ordering of speech. *Psychological Review*, 1970, **77**, 182–196.

Martin, J. G. Rhythm vs. grouping in memory for digits. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1969. (a)

Martin, J. G. Temporal structure and the perception of speech. Paper presented at the meeting of the Midwestern Psychological Association Chicago, May 1969. (b)

Martin, J. G. Rhythm-induced judgments of word stress in sentences. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 627–633. (a)

Martin, J. G. The rhythmic structure of speech, music and other meaningful sounds. Paper presented at the meeting of the Psychonomic Society, San Antonio, November 1970. (b)

Mattingly, I. G. Synthesis by rule as a tool for phonological research. *Language and Speech*, 1971, **14**, 47–56.

Mattingly, I. G. Synthesis by rule of prosodic features. *Language and Speech*, 1966, **9**, 1.

McNeill, D. Sentences as biological processes. Paper presented to the CNRS Conference on Psycholinguistics, 1971.

Miller, G. A. Decision units in the perception of speech. *IRE Transactions on Information Theory*, 1962, **8**, 81–83.

Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.

Milner, P. *Physiological psychology.* New York: Holt, Rinehart & Winston, 1970.

Neisser, U. *Cognitive psychology.* New York: Century, 1967.

Neisser, U., Hoenig, Y. J., & Goldstein, E. Perceptual organization in the prefix effect. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 424–429.

Newman, S. E. On the stress system of English. *Word*, 1946, **2**, 171–187.

Pierce, J. R. Whither speech recognition? *Journal of the Acoustical Society of America*, 1969, **46**, 1049–1051.

Pike, K. L. *Intonation of American English.* Ann Arbor: University of Michigan, 1945.

Preusser, D., Garner, W. R., & Gottwald, R. L. Perceptual organization of two-element temporal patterns as a function of their component one-element patterns. *American Journal of Psychology*, 1970, **83**, 151–170.

Rabiner, L. R. A model for synthesizing speech by rule. *IEEE Transactions on Audio and Electroacoustics*, 1969, **AU-17**, 7–13.

Rabiner, L. R., Levitt, H., & Rosenberg, A. E. Investigation of stress patterns for speech synthesis by rule. *Journal of the Acoustical Society of America*, 1969, **45**, 92–101.

Restle, F. Theory of serial pattern learning: Structural trees. *Psychological Review*, 1970, **77**, 481–495.

Restle, F. Serial patterns: The role of phrasing. *Journal of Experimental Psychology*, 1972, **92**, 385–390.

Restle, F., & Brown, E. Organization of serial pattern learning. In G. H. Bower (Ed.), *Psychology of learning and motivation.* New York: Academic Press, 1970

Royer, F. L., & Garner, W. R. Response uncertainty and perceptual difficulty of auditory temporal patterns. *Perception and Psychophysics*, 1966, **1**, 41–47.

Royer, F. L., & Garner, W. R. Perceptual organization of nine-element auditory temporal patterns. *Perception and Psychophysics*, 1970, **7**, 115–120.

Schane, S. A. *French phonology and morphology.* Cambridge: M.I.T. Press, 1968.

Sturges, P. T., & Martin, J. G. Rhythmic structure in auditory temporal patterns. Paper presented at the meeting of the Psychonomic Society St. Louis, November 1971.

Tatham, M. A. A. A speech production model for synthesis-by-rule. (Working Papers in Linguistics, No. 6) Columbus: Ohio State University, 1970.

Teranishi, R., & Umeda, N. Use of pronouncing dictionary in speech synthesis experiments. *Reports of the Sixth International Conference on Acoustics*, 1968, **2**, 155–158.

Trager, G. L., & Smith, H. L. *Outline of English structure.* (Studies in Linguistics. Occasional Paper No. 3) Norman, Okla.: Battenburg Press, 1951.

Vanderslice, R. Synthetic elocution. (Working Papers in Phonetics, No. 8) Los Angeles: University of California, 1968.

Vanderslice, R. Occam's razor and the so-called stress cycle. *Language Sciences*, 1970, **13**, 9–15.

Weir, R. H. Some questions on the child's learning of phonology. In F. Smith & G. A. Miller (Eds.), *The genesis of language.* Cambridge: M.I.T. Press, 1966.

Wickelgren, W. A. Context sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, 1969, **76**, 1–15.

Wiersma, C. A. G. (Ed.) *Invertebrate nervous systems.* Chicago: University of Chicago Press, 1967.

Woodrow, H. S. A quantitative study of rhythm. *Archives of Psychology*, 1909, **18**, No. 1.

Woodrow, H. S. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

(Received March 23, 1972)

Psychological Review
1972, Vol. 79, No. 6, 510–524

PERCEPTION OF ROTATION IN DEPTH:

A PROCESS MODEL[1]

MYRON L. BRAUNSTEIN[2]

University of California, Irvine

A theory is proposed which accounts for both veridical and illusory perception of direction of rotation of rectangles and trapezoids on the basis of the same heuristic perceptual processing rules and the same sources of information. The change in angle between horizontal and vertical contours is held to be the principal indicator of direction of rotation, with the difference in acceleration of the vertical sides playing a secondary role. A process model is presented in which these sources are combined hierarchically. Predictions of the model match the quantitative results of two studies used in its development and the qualitative results of most other investigations. The role of heuristic processes in the model and the relationship of the model to other theories of perceived rotary motion are discussed.

The role of complex motions of contours in the optic array has gained increasing attention in the study of depth perception over the past two decades. The most widely studied stimulus transformation has been the rotation of an object or of a pattern about a vertical axis perpendicular to the line of sight. There have been two lines of research on the perception of this transformation which have only recently become fully merged. The first line of research has been concerned with the conditions under which the retinal projections of various objects and patterns undergoing this transformation serve as stimuli for the perception of motion in three-dimensional space. The studies conducted have attempted to isolate the effects of transformations by eliminating disparity, convergence, and accommodation as potential indicators of differential distance of parts of the rotating object. This was accomplished in the earlier studies by displaying shadows of rotating objects (Gibson & Gibson, 1957; Miles, 1931; Wallach & O'Connell, 1953; White & Mueser, 1960) and in later studies through the use of computer animation (Braunstein, 1962a, 1966; Green, 1961; Johansson, 1964). A general finding of this line of research is that the accuracy of direction of rotation judgments tends to be high when the shadow or computer-generated projection is a polar projection but drops to chance levels when a parallel projection is displayed (Braunstein, 1962b, 1966; Gibson, 1957). The latter result is not surprising as a parallel projection of an object rotating about an axis perpendicular to the line of sight is identical for clockwise and counterclockwise directions of rotation.[3]

---

[1] This paper is based on research supported by National Science Foundation Grants GB 2335, GB 3923, GB 8251, and GB 28704X. The author is indebted to John W. Payne for his assistance in the preparation of this paper and in the research on which it is based and to John S. Brown for his critical reading of an earlier version of this paper.

[2] Requests for reprints should be sent to Myron L. Braunstein, School of Social Sciences, University of California, Irvine, California 92664.

[3] Polar projections are produced when the light reflected by an object is projected onto a surface through a point, where this projection point is at a finite distance from the object. In direct vision, the surface is the retina and the point is the focal point of the eye. Direct vision always involves polar projections. As the distance of the observed object from the eye increases, the projection approaches a parallel projection. When an object is so distant that it must be observed through a telescopic lens, the projection is essentially parallel. When shadow projections are produced with a point light source the distance of the light source from the shadow-caster, relative to the extent in depth of the shadow-caster, determines the type of projection. A true parallel projection would require an infinitely distant light source, but sufficiently great distances can be used in the laboratory to approximate parallel projections. Polar projections approximating direct vision can be produced by the shadow method by equating the light source to screen distance with the subject to screen distance. The computer method involves the same geometric considerations as the shadow method. The coordinates

Gibson (1957) explains the perception of rigid motion in displays of shadows of objects undergoing continuous transformations on the basis of a direct sensitivity of the eye to specific transformations. When the optical motions induced at the eye by the shadows of transformations are the same as the motions which would occur in direct vision, the shadows serve as an effective stimulus for depth perception. Johansson (1970) has expressed disagreement with the implication that the perception of rigid transformations in three-dimensional space can be understood simply by studying the geometric information in the stimulus. He states that empirical and theoretical work should be directed at discovering the rules for stimulus data treatment which characterize the visual system. The strategy followed in the development of the model presented below is in accord with Johansson's proposal.

The second line of research has been concerned mainly with the perception of the direction of rotation of rectangles and trapezoids. The objects have usually been viewed directly. Depth information from disparity and convergence has been controlled through the use of monocular viewing. Viewing distances have usually been used at which accommodation is ineffective. The principal finding of this research is that under the same viewing conditions, accurate reports of direction of rotation are obtained for rectangles but reports of oscillation are obtained for trapezoids. This is the well-known rotating trapezoid illusion (Ames, 1951).

A number of explanations have been offered for the apparent oscillation of a rotating trapezoid. Ames (1951) held the decrease in the overall horizontal projection of the form to be the basis for perceived rotary motion. He further proposed that subjects unconsciously assumed the trapezoid to be a slanted rectangle because of their past experience with the retinal projections of rectangular windows viewed from various angles. This assumption would cause the shorter vertical side to appear tilted away from the subject at all times. When the shorter side was actually closer, the orientation of the trapezoid would be misperceived. The combination of perceived rotary motion and misperceived orientation would result in misperceived direction of rotation. Pastore (1952) and Graham (1963) offered explanations similar to that proposed by Ames, except for their handling of misperceived initial slant. Pastore proposed that the initial misperception of slant was due to central processes of the kind advanced by Gestalt psychologists. Graham concluded that it was linear perspective which caused the longer vertical side to appear closer, regardless of its actual position. Day and Power (1965) proposed a theory based on an identity of retinal projections for the two directions of motion. Hershberger (1967) has noted that this identity holds only for parallel projections. The Day and Power explanation was also disputed by Cook, Mefferd, and Wieland (1967). They concluded that apparent reversals of rotary motion are due to apparent reversals of orientation (perspective reversals) in depth which "occur quite automatically and abruptly, even in the presence of valid depth cues [p. 701]." They do not, however, explain the basis for the reversals.

There is a general problem with these explanations cited above for the perception or misperception of rotary motion: These explanations fail to take into account the interaction between changes in the shape of the projection of a rotating form which are due to characteristics of its original shape and changes which are due to the distance between the form and the projection point (see Footnote 3). Specifically, they fail to account for the most general findings of the studies previously cited: With a sufficiently distant projection point the direction of rotation of any form may be misperceived, with certain forms giving rise to systematic misperceptions. With closer projection points, the direction of rotation of some forms will

---

of the projection point, analogous to the light source, and the projection plane, analogous to the screen, are used to compute a projection which is then displayed on a cathode-ray-tube and usually photographed to form the frame of a motion picture. As the computer technique is not subject to the physical limitations of light source to object distance, it can be used to produce true parallel projections.

be correctly perceived while that of other forms will still be misperceived. With a close enough projection point, the direction of rotation of any form will be correctly perceived.

These findings were probably not incorporated into a general theory of the perception of rotary motion because they derived from separate lines of research employing separate methodologies. The shadow projection method had not been used in comparisons of accuracy of direction judgments for rectangles and trapezoids. Studies comparing these two figures had not included actual or simulated variations in viewing distance (although Zegers, 1964, studied these two variables in separate experiments in the same series). This lack of information on the interaction of two principal stimulus variables affecting the transformation sequence projected at the eye during observation of rotation in depth demonstrated a need for additional research before an adequate model could be developed. The two studies summarized below (Braunstein, 1971; Braunstein & Payne, 1968), represent the beginnings of this line of research. They form the empirical base for the model presented in the section which follows.

EXPERIMENTAL DATA

The purpose of the first study was to determine the effects of projection distance on judgments of direction of rotation for two-dimensional forms. The stimuli were computer-generated motion picture sequences representing rectangles, trapezoids, circles, and ellipses undergoing 360° rotations about a vertical axis. Projections of these transformations were computed with projection points located at infinity and at distances from the axis of rotation equal to 4.5, 2.5, 1.83, and 1.5 times the width of the figure. The first projection point location yields a parallel projection; the other four yield polar projections. Accuracy of direction judgments was low for all forms displayed through parallel projections, but the type of error varied with the figure. The direction of rotation of the rectangle was accurately reported for the entire 360° cycle in one-fourth of the observations. The direction was inaccurately reported for the entire cycle in approximately one-fourth of the observations. Reversal of direction during a 360° cycle was reported in the remaining half of the observations. In contrast, oscillation was reported in every 360° cycle when the trapezoid was displayed with a parallel projection. With polar projections, accuracy increased with decreasing projection point distance for all forms, but increased most rapidly for the rectangle. Sequences were also produced in which only the vertical or only the horizontal coordinates of each projected contour were affected by a polar projection based on one of the four finite projection distances, with the other coordinates subjected to a parallel projection. These special projections were used in an attempt to separate the effects of changes in each of the dimensions of the projected image on judgments of direction of rotation. Figure 1 shows views of a rotated rectangle displayed with an ordinary polar projection, with vertical changes only based on a polar projection and with horizontal changes only based on a polar projection. Accuracy of direction judgments for rotating rectangles and trapezoids increased about as rapidly with decreasing projection point distance when this distance affected only the vertical dimensions of the projected image as when it affected both dimensions. When distance changes affected the horizontal dimension only, accuracy exceeded chance levels but remained low for all projection distances. It was concluded that the changes in the image projected by a rotating rectangle or trapezoid which most affect accuracy of direction of rotation judgments are in the dimension parallel to the axis of rotation (vertical in that study). Changes in the other dimension appeared to have a secondary effect. This implies that direction judgments are affected by at least one primary factor related to changes in the dimension parallel to the axis of rotation and at least one secondary factor related to changes in the perpendicular dimension of the projected image. (Support for this implication is also found in a study by Jansson & Borjesson, 1969,

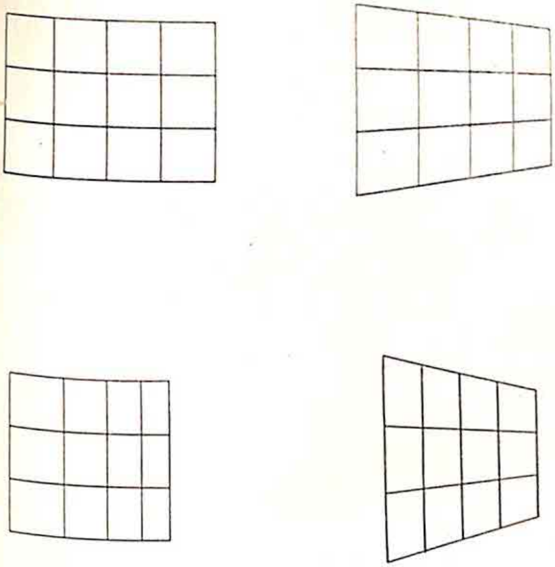


FIG. 1. Projections of a rectangle (upper left) rotated about a vertical axis and displayed with a polar projection (upper right), with horizontal changes only affected by a polar projection (lower left) and with vertical changes only affected by a polar projection (lower right).

using a different method of isolating horizontal and vertical changes in polar projections.)

The second study attempted to further specify the sources of information used in direction of rotation judgments through the use of a series of forms which systematically combined the geometrical features of rectangles and trapezoids. These features were encompassed in three factors: (*a*) angles between horizontal and vertical contours; (*b*) extents of vertical contours; and (*c*) position of the axis of rotation relative to the vertical contours (see Figure 2). The axis of rotation was either at the horizontal center of the form (symmetric) or displaced to the right (asymmetric). (In the Ames trapezoid and in those studied by Braunstein and Payne, 1968, the axis of rotation was closer to the shorter vertical side.) Each form was displayed undergoing one complete rotation about a vertical axis. Each display was produced with one of five distances between the projection point and the axis of rotation. Accuracy of direction judgments increased with decreased projection distance for each form. Accuracy was highest for the forms most like rectangles, lowest for those most like trapezoids, and intermediate for the forms which combined features of rectangles and trapezoids. For the intermediate forms, accuracy was ordered primarily by angle relationships, secondarily by side relationships, and affected only slightly by the position of the axis of rotation. These results, together with those of the Braunstein and Payne study, indicated that changes in the angles between horizontal and vertical contours in the projected image constitute the principal source of information used by subjects in direction of rotation judgments. Changes in these angles are affected by direction of rotation only in polar projections and reflect changes in the vertical positions of the projected contours.

There appears to be a secondary source of information related to changes in the horizontal positions of the projected contours. The projections of both vertical sides increase in velocity as the form leaves the frontal position. These velocity changes are affected by direction of rotation only in polar projections. It is hypothesized that subjects tend to judge the side which is accelerating more rapidly to be approaching. The use of acceleration differences as the secondary factor in the model is a somewhat arbitrary choice. It is based on the similar-

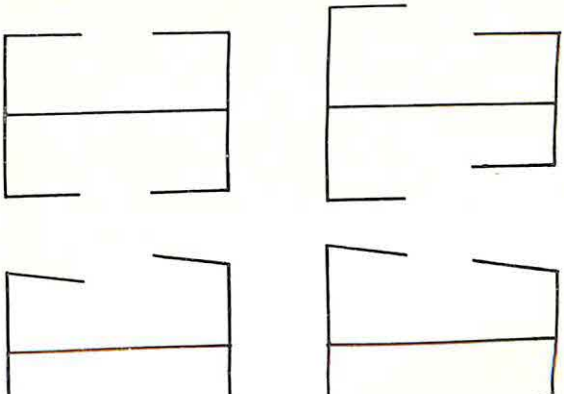


FIG. 2. An incomplete rectangle (upper left) and trapezoid (lower right) and forms with the angles of a rectangle and the vertical sides of a trapezoid (upper right) and the sides of a rectangle and the angles of a trapezoid (lower left).

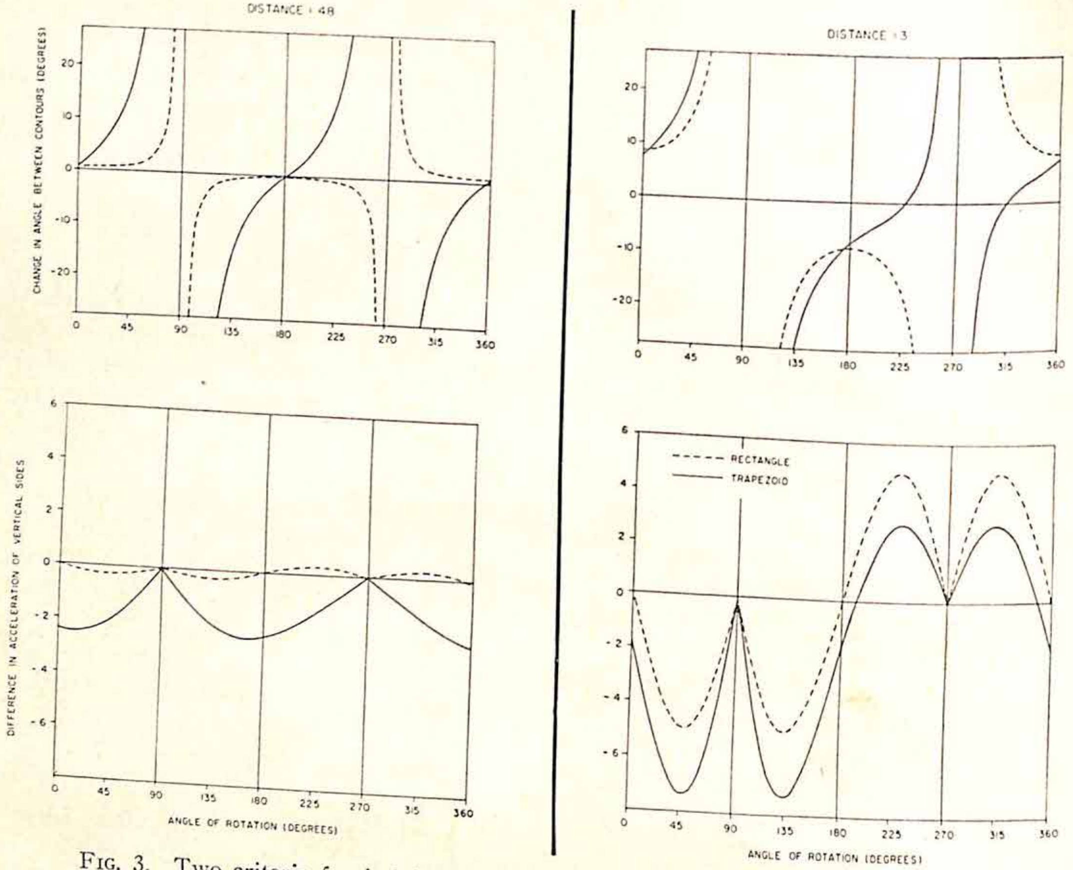


FIG. 3. Two criteria for judging direction of rotation for a rectangle and a trapezoid undergoing counterclockwise rotation, shown for two viewing distances (in arbitrary units; form width was 1.2 units).

ity of the relationship between acceleration differences and direction judgments to that found between angle changes and direction judgments. In both cases, direction judgments appear to be related to the sign of the criterion variable.

The two hypothesized criteria for judgments of direction of rotation are illustrated in Figure 3. At the greater viewing distance (upper left), the contour angle change function has the same sign for both the rectangle and the trapezoid, between 0° and 180°. The signs are opposite for the two forms between 180° and 360°. A rule based on the sign of this function which would produce correct direction judgments for rectangles would produce judgments of oscillation for trapezoids. At the closer viewing distance (upper right), the sign remains the same for both forms for a considerable interval (over 45°) after the forms return to the frontal position (180°). A rule based on the sign of this function and applied within that interval would produce the same direction judgments for both forms.

The lower graphs illustrate an analogous situation for the hypothesized secondary source of information: the difference in acceleration of the vertical sides. At the greater viewing distance (lower left), the sign of this function is the same for both forms between 0° and 180° and opposite for the two forms between 180° and 360°. At the closer viewing distance (lower right), the sign is again the same for both forms between 0° and 180°. The two functions take on opposite signs at 180°, but the sign is again the same for both forms after 192°. A rule based on the sign of this function and applied at least 12° after the form leaves the frontal position would produce the same direction judgments for both forms. Either criterion, applied in an interval between 12° and 45° after the form leaves the frontal position, could account for (*a*) correct perception of direction of rotation for rectangles at the greater viewing distance, (*b*) perception of oscillation for trapezoids at the greater distance, and (*c*) correct perception of direction for both forms at the closer distance.

If these graphs were presented for parallel projections, both the contour change and acceleration difference functions for the rectangle would appear as horizontal lines through $Y = 0$. Neither criterion is usable for direction judgments in this case, and the model presented below matches the experimental data with a random process. For the trapezoid, these functions would be similar to those shown for the distance of 48. Subjects are assumed to use these two criteria in judging direction of rotation, even for parallel projections of a rotating trapezoid, although these functions are unrelated to direction of rotation for such projections. They are determined entirely by the shape of the trapezoid and the position of the axis of rotation. The angle change function, for parallel projections, deviates from a horizontal line as the angles in the original form deviate from right angles. The acceleration difference function deviates from a horizontal line as the vertical sides in the original form deviate from equidistance from the axis of rotation.

A MODEL OF THE PERCEPTION OF ROTARY MOTION

The following model is based on the two processing rules suggested in the previous section. It is intended to account for the responses obtained in the two studies just discussed and to be consistent with general findings on the perception of direction of rotation for rectangles and trapezoids. The presentation of these rules in the form of a process model should facilitate generalization to additional two- and three-dimensional stimulus materials and to other experimental conditions.

Before judging direction of rotation, the subject must decide (*a*) whether he is observing a rigid form, and (*b*) whether it is rotating in three-dimensional space.

> These two related judgments are not handled in the present model, but appear to involve a comparison of the relative changes in horizontal and vertical positions of projected points over time to a cosine function.

If the form is judged to be rigid and rotating, its direction is judged in the following manner:

> One of the vertical sides is considered. (This choice is arbitary in the present model.)
>
> A decision is made as to whether that side is approaching or receding: If changes in the angles between that side and the projection of the horizontal sides are detectable, then (*a*) if the angle is increasing, the side is judged to be receding, or (*b*) if the angle is decreasing the side is judged to be approaching.
>
> If changes in the angles are not detectable, then if changes in the relative acceleration of the vertical sides are detectable, the side which is accelerating more rapidly is judged to be approaching.
>
> If neither change is detectable, the judgment of which side is approaching is made on some other basis which will tentatively be approximated by a random choice.

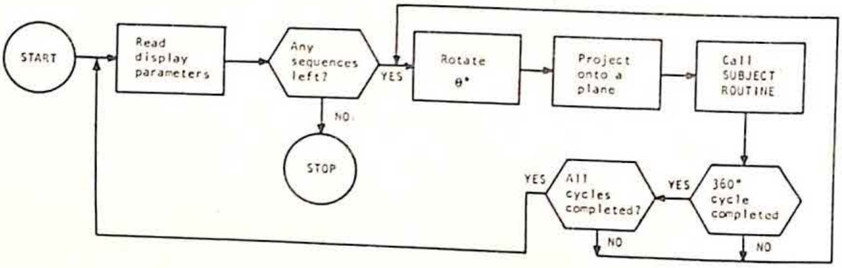
If the side is judged to be approaching, then:

> if it is to the right of the axis of rotation a clockwise direction is judged.
>
> Otherwise, a counterclockwise direction is judged.

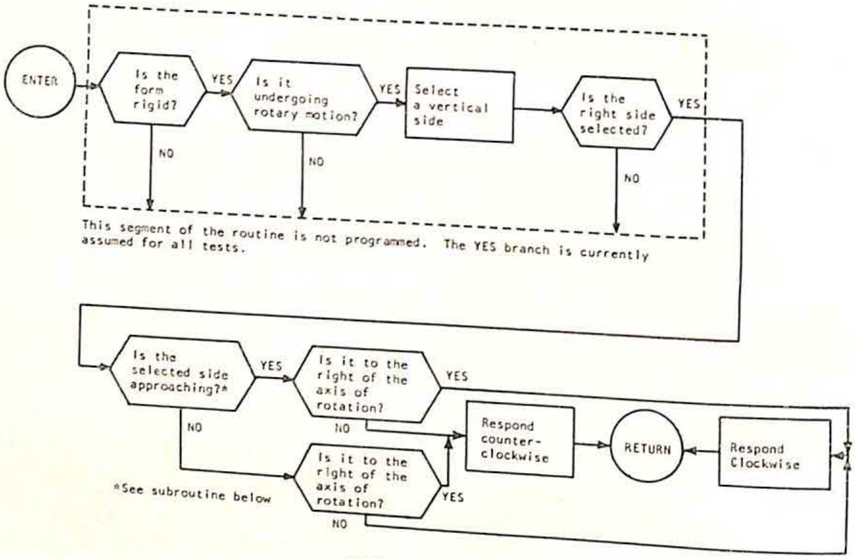
If the side is judged to be receding, then:

> if it is to the right of the axis of rotation a counterclockwise direction is judged.
>
> Otherwise, a clockwise direction is judged.

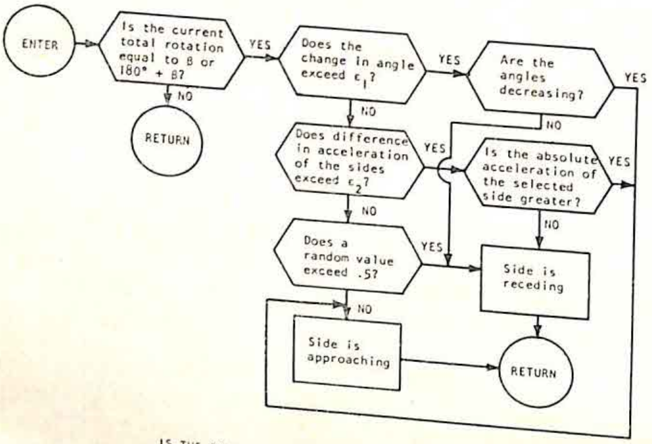
Figure 4 presents the model in flow chart form. The "Experimenter Routine" is not part of the model but is included to show the nature of the stimulus input to the "Subject Model." The information provided to the Subject Model is the same as that used to plot frames of computer-generated motion pictures for use in experiments with human subjects (Braunstein, 1962a, 1966, 1971; Braunstein & Payne, 1968).



EXPERIMENTER ROUTINE


ENTER
Is the form rigid?
YES
NO
Is it undergoing rotary motion?
YES
NO
Select a vertical side
Is the right side selected?
YES
NO
This segment of the routine is not programmed. The YES branch is currently assumed for all tests.
Is the selected side approaching?*
YES
NO
Is it to the right of the axis of rotation?
YES
NO
Is it to the right of the axis of rotation?
YES
NO
Respond counter-clockwise
RETURN
Respond Clockwise
*See subroutine below

SUBJECT ROUTINE


ENTER
Is the current total rotation equal to β or 180° + β?
YES
NO
RETURN
Does the change in angle exceed ϵ_1?
YES
NO
Are the angles decreasing?
YES
NO
Does difference in acceleration of the sides exceed ϵ_2?
YES
NO
Is the absolute acceleration of the selected side greater?
YES
NO
Does a random value exceed .5?
YES
NO
Side is receding
Side is approaching
RETURN

IS THE SELECTED SIDE APPROACHING?

FIG. 4. A model for the perception of rotating figures.

The first three decisions in the Subject Model are not fully specified in the present version of that model. Data on the processes involved in these decisions are inadequate at this time. The present version uses the "yes" branch in these three tests. This does not present a problem when the model's application is limited to the perception of rotating rectangles and trapezoids. The perception of a rigid form undergoing rotary (including oscillatory) motion occurs almost without exception for these stimuli. The selected side makes little difference. These steps would have to be specified before the model could be applied to a wider range of stimuli.

The central decision in the model, "Is the selected side approaching?," is handled in a subroutine. In the present version of the model this question is asked at only two positions during each 360° rotation. This implies that changes in perceived direction of rotation occur within a small interval ($\beta°$) after the form leaves the frontal position (0° or 180°). This result has been obtained in several studies, with a variety of forms (Cook, Mefferd, & Wieland, 1967; Day & Power, 1965; Graham & Gillam, 1970). If the rotation sequence is at $\beta°$ or at $180 + \beta°$, the first decision criterion is checked. (Testing direction at the end of the $\beta°$ interval, rather than throughout the interval, is an approximation which has no effect on the performance of the model.) If the absolute change in the angle between the selected vertical side and an adjacent segments exceeds a subject sensitivity parameter, the decision of whether the side is approaching is made on the basis of that criterion. The second criterion is considered if the first criterion does not exceed the sensitivity parameter. If the difference in acceleration of the two vertical sides does not exceed a second sensitivity parameter, the decision is made by a random process. The use of a random process, it is hoped, is a temporary expedient which can eventually be replaced with additional decision criteria.

Once the decision of whether the selected side is approaching is made by the subroutine, the judgment of direction of rotation by the Subject Model is straightforward. It is assumed that the subject can judge the location of the axis of rotation (which is obvious at 90° and 270°) and can easily determine whether the projection of the selected side is moving to the left or to the right across his field of view.

The model handles individual differences by using a range of subject sensitivity parameters. At present, it is designed to compute averages for simulated subjects, predicting group data. If an independent measure of the sensitivity parameters could be obtained for individual subjects, and if these parameters represent a subject characteristic which is stable over time, the model could be used to predict the judgments of individual subjects.

EVALUATION OF THE MODEL

In a first test of the model, a set of subject sensitivity parameters and a position sampling parameter were selected which would enable the model to approximate the quantitative results of Braunstein and Payne (1968) and unpublished detailed results summarized by Braunstein (1971). The model results were based on 100 simulated subjects. Each subject was assigned to one of three levels of sensitivity to contour angle changes for rectangles and to one of three different levels of sensitivity to contour angle changes for trapezoids. The use of two sets of sensitivity parameters for angle changes reflects the hypothesis (Braunstein & Payne, 1968) that subjects are more sensitive to deviations from right angles than to deviations from acute or obtuse angles. Each subject was assigned to one of two levels of sensitivity to acceleration differences. These levels were the same for both rectangles and trapezoids. The deviation from the frontal plane at which the model made its direction judgments was 36°.

Table 1 compares the observed results with those generated by the model. The model accounts for the quantitative differences in the results obtained with rectangles and trapezoids in both studies. Results matched by the model include (*a*) the finding of 50% oscillations per cycle for rectangles and 100% for trapezoids with parallel projections; (*b*) the more rapid increase to complete accuracy for rectangles than for trapezoids with decreasing projection distance; and (*c*) the slight increase in accuracy found at intermediate projection distances when the axis of rotation was at the center of the form rather than displaced to one side.

TABLE 1

COMPARISONS OF RESPONSE PROPORTIONS GENERATED BY THE MODEL TO OBSERVED PROPORTIONS

Figure	Perspective ratio[a]	Model			Observed		
		Correct	Incorrect	Oscillation	Correct	Incorrect	Oscillation
Symmetric rectangles	1.00[b]	.18	.30	.52	.25	.29	.46
	1.025[c]	.57	.22	.21	.54	.04	.42
	1.051[c]	.61	.17	.22	.58	.08	.33
	1.11[c]	.84	.06	.10	.83	.08	.08
	1.22[c]	.93	.03	.04	1.00	0	0
	1.25[b]	1.00	0	0	.98	.02	0
	1.50[b]	1.00	0	0	1.00	0	0
	1.50[c]	1.00	0	0	1.00	0	0
	1.75[b]	1.00	0	0	1.00	0	0
	2.0[b]	1.00	0	0	.99	0	.01
Asymmetric rectangles	1.025[c]	.57	.10	.33	.58	.08	.33
	1.051[c]	.58	.11	.31	.71	.04	.25
	1.11[c]	.86	.02	.12	.83	0	.17
	1.22[c]	.84	.02	.14	.87	.08	.04
	1.5[c]	1.00	0	0	.92	.04	.04
Symmetric trapezoids	1.025[c]	0	0	1.00	.04	.04	.92
	1.051[c]	.07	0	.93	.17	0	.83
	1.11[c]	.09	0	.91	.08	0	.92
	1.22[c]	.39	0	.61	.46	0	.54
	1.50[c]	.93	0	.07	.92	0	.08
Asymmetric trapezoids	1.00[b]	.03	.03	.94	0	0	1.00
	1.025[c]	0	0	1.00	0	0	1.00
	1.051[c]	.07	0	.93	.08	0	.92
	1.11[c]	.09	0	.91	.04	0	.96
	1.22[b]	.46	0	.54	.31	.03	.66
	1.22[c]	.21	0	.79	.25	0	.75
	1.43[b]	.98	0	.02	.84	.02	.14
	1.50[c]	.78	0	.22	.79	0	.21
	1.64[b]	1.00	0	0	.96	.02	.02
	1.84[b]	1.00	0	0	1.00	0	0

[a] The perspective ratio is the ratio of the sum of the distance of the projection point to the axis of rotation and half the width of the form to the difference between that distance and half the width of the form. A ratio of 1.0 represents a parallel projection (see Braunstein, 1971).

[b] Observed proportions from Braunstein and Payne (1968). The perspective ratios for the trapezoids in the table are based on the width of the trapezoids and therefore differ from those in the original article which were based on the width of the rectangle used to generate the trapezoids.

[c] Observed proportions from Braunstein (1971).

(Compare, for example, the symmetric and asymmetric rectangles at the 1.22 perspective ratio or the symmetric and asymmetric trapezoids at the 1.50 perspective ratio.)

As a further test of the model, an attempt was made to match the results of five other studies (Cook, Mefferd, & Wieland, 1967; Day & Power, 1963; Murch, 1970; Power, 1967; Zegers, 1964), without altering the values of any of the parameters. Each of these studies differed from the first two to which the model was compared in three ways: (*a*) Subjects observed rotating objects directly, rather than two-dimensional projections of rotating forms. (*b*) A trial consisted of 20–30 revolutions, as compared to a single revolution in the first two studies. (*c*) Subjects signaled when they wished to report a reversal of direction, rather than continuously indicating one direction or another. Zegers (1964) presented the most extensive set of quantitative results. Figure 5 compares his results with the predictions of the model for an experiment in

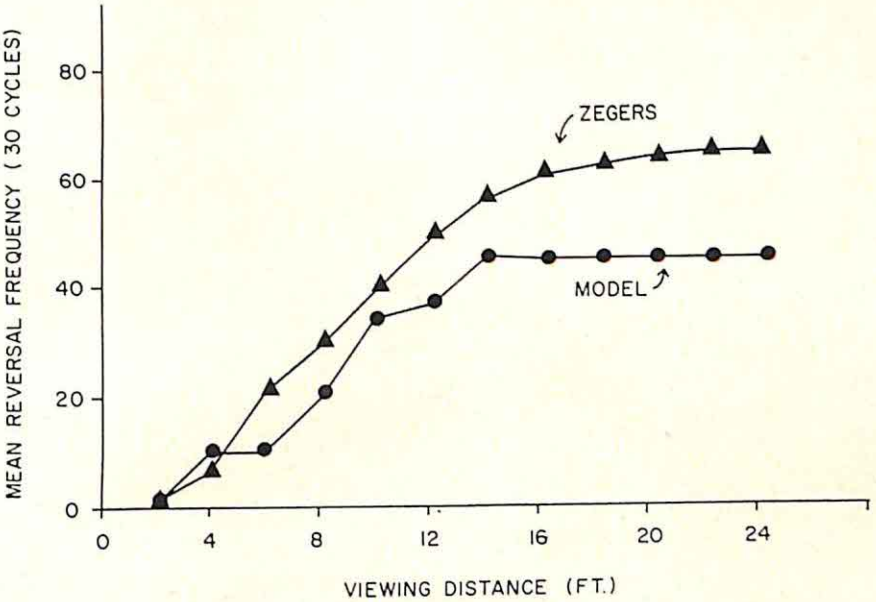


FIG. 5. Comparison of the model to results of Zegers (1964) for variations in viewing distance.

which reversal frequency was measured as a function of viewing distance. The model follows the observed trend but underestimates reversal frequencies at most distances. This can be attributed to a difference between the rotation speed used by Zegers (25 revolutions per minute [rpm]) and that used in the studies on which the parameters of the model are based (8 rpm). Zegers found differences in reversal frequency of this magnitude as a result of differences in rotation speeds in another experiment reported in the same paper. Figure 6 compares Zegers' results with the predictions of the model when the length of one of the vertical sides of the form is varied. The rotation speed used by Zegers in the part of his experiment illustrated was 5 rpm. The observed and model results are similar in general trend and in quantitative level. A more detailed comparison would probably not be meaningful as Zegers' points are each based on data from only three subjects, viewing each rotating object for one 30-revolution trial.

The model cannot match the quantitative results of Cook, Mefferd, and Wieland (1967). These results are quite different from those of other investigators using similar stimulus materials and viewing distances (e.g., Day & Power, 1963; Murch, 1970; Power, 1967; Zegers, 1964). At a rotation speed of 5 rpm, for example, they obtained 12 times the reversal frequency for a rectangle and twice the frequency for a trapezoid as did Zegers (1964) using comparable figure dimensions and viewing distance. Unlike other investigators, Cook et al. failed to find a significant difference between rectangles and trapezoids. These discrepancies may be related to the unusual subject population in the Cook study. Their subjects

> had had considerable experience in making perceptual judgments but were naive with respect to the purposes of these experiments. Motivated *O*s who were experienced in making careful perceptual judgments and reporting them in detail were deemed desirable for several reasons. The present experiments required a number of very fine perceptual discriminations which were not easily made by naive *O*s. The probability of obtaining reliable and valid judgments was therefore maximized with practiced *O*s. Moreover, many perceptual phenomena are overlooked by naive *O*s unless a period of specific training and instruction is given [p. 694].

Predictions of the results of other recent studies preserve the ordinal relationships in

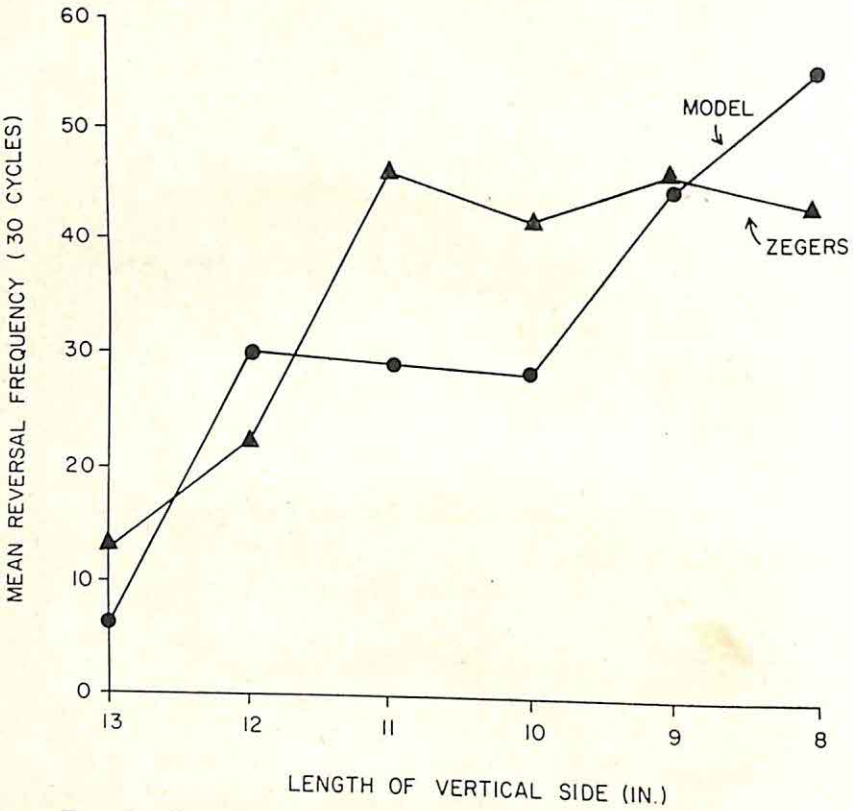


FIG. 6. Comparison of the model to results of Zegers (1964) for variations in trapezoid dimensions.

the obtained data, although they deviate from the quantitative values reported. For 20 revolutions, mean apparent reversals were obtained of 0 for a rectangle (Day & Power, 1963), 2.9 (Power, 1967), and 3.9 (Murch, 1970) for squares, and 34 (Day & Power, 1963), and 36 (Murch, 1970) for trapezoids. The model predicts means of 2, 4.7, 6.6, 15, and 22, respectively, for the form dimensions and viewing distances used in those studies.

In summary, the model can match the detailed quantitative results of the two studies used to develop it and, without parameter adjustment, can approximate the quantitative findings of the Zegers (1964) studies and the qualitative results of other experimenters. The quantitative results obtained with the model depend on subject sensitivity parameters which should be independently measurable. When more data is available on the effects of experimental procedures on these parameters, it should be possible to make meaningful parameter adjustments. This should enable the model to fit the quantitative results of a wider variety of experiments. At present, the overall agreement between the ordinal predictions of the model and results obtained in widely differing experimental situations can be regarded as support for this form of model.

DISCUSSION

Heuristic Processes in Depth Perception

The two principal processing rules incorporated in the model are examples of heuristic processes. These are problem-solving methods which tend to produce efficient solutions to difficult problems by restricting the search through the space of possible solutions, on the basis of some evaluation of the structure of the problem. The same restriction on search which increases efficiency when a solution is reached by means of a heuristic process sometimes results in failure to reach the desired solution. The definition of a desired solution and the criteria for efficiency vary with the particular problem. In theorem-proving, for example, the desired solution is a valid proof. Heuristic processes are used to minimize the number of permutations of axioms which must be tried. In chess, the desired solution is the selection of a move which maximizes the player's likelihood of not losing the game. Heuristic methods are used to reduce the depth of the analysis required to select the move. Heuristic methods are usually contrasted with algorithms which would guarantee the desired solution if fully executed, but which may be too slow or otherwise inefficient to be useful in practice. Trying all permutations of axioms to prove a theorem or following the implications of every possible chess move to the end of the game are examples of such algorithms.

In depth perception, heuristic methods are used to solve a continuous problem for the human observer: the problem of transforming information from the optic arrays in the two eyes and adjustments of the muscles involved in convergence and accommodation into hypotheses concerning the external environment. The hypotheses represent the observer's solutions to the perceptual problem. Veridical perception would appear to be an appropriate definition of a desired solution. Perceptual processes which regularly misinformed an animal about his environment would not be expected to be found in species, or in individuals, which have survived (Gibson, 1966, Ch. 1).

Heuristic processes occasionally fail to achieve veridical perceptions even when sufficient information is available in the optic array. These failures do not represent errors in the application of processing rules. Nonveridical perception occurs when these rules are correctly applied to stimulus situations which are not in accord with the assumptions about the visual environment on which the rules are based. When veridical perception is achieved, the heuristic processes used are probably more efficient than alternative processes which would take full advantage of the information available in the optic array.

There are at least two plausible criteria for the efficiency of a perceptual process. The process should be capable of achieving certain critical judgments very rapidly. Judgments about whether a form, even if unidentified, is approaching or receding, may fall into this category. A slower process which would always lead to a correct judgment might be of less value in helping an observer select appropriate responses than a more rapid process which might lead to an incorrect judgment for very unusual stimulus conditions. A useful perceptual process should also be capable of handling degraded information. Processes which always lead to correct judgments under ideal conditions, but which do not work at all under degraded conditions, may not be as useful as processes which usually lead to correct judgments under ideal conditions and often do so under degraded conditions.

In the present model, the decision rules based on changes in contour angles and differences in acceleration are believed to reflect the principal heuristic processes used by human subjects. Either of these processes can produce accurate direction judgments for polar projections of rotating rectangles. The first process produces inaccurate judgments for trapezoids at low perspective levels when the vertical side enclosed by acute angles is receding. The second process produces inaccurate judgments at low perspective levels for figures with vertical sides which are not equidistant from the axis of rotation when the vertical side further from that axis is receding. There is another process which, if used, would lead to accurate judgments of direction of rotation for both rectangles and trapezoids at all polar perspective levels. This would simply require that a side be judged as approaching if the visual angle it subtends is increasing. Subjects do not appear to make judgments of direction of rotation on that basis. The arguments presented above would suggest that subjects might not use that process because it is based on changes which are more difficult to detect or are more susceptible to degradation than are contour angle changes and acceleration comparisons. The type of evidence which would be required to confirm this suggestion is not currently available, so it must remain in the realm of speculation for now.

Relationships to Other Theories

There is a fundamental difference between the present theory and others which have offered explanations for the mispreception of rotary motion in depth. Each of the explanations previously cited for the rotating trapezoid illusion holds that observers fail to use veridical motion information because this information is weak, absent, or dominated by information from a nonveridical source. In the present theory the misperception of rotary motion does not involve a conflict between veridical and nonveridical sources of information or the switching from a veridical to a nonveridical source. Instead, the same sources of information and the same processing rules which result in veridical direction judgments for some forms under some viewing conditions have been shown to result in nonveridical judgments for other forms under other conditions. These sources are related to dynamic changes in the projection of the rotating form rather than to the perception of static characteristics of the form. It seems unlikely that direction judgments would be based on perceptions of static properties of a form which is shown undergoing a continuous transformation. On the contrary, dynamic changes have been shown to be considerably more effective than static properties in determining slant judgments (Braunstein, 1968; Gibson & Gibson, 1957). As the present explanation of nonveridical direction judgments is not based on misperceived initial slant, there is no need to address the question of whether this misperception is due to past experience with windows (Ames, 1951), Gestalt factors (Pastore, 1952), linear perspective (Graham, 1963), or false shape constancy (Gibson & Gibson, 1957).

An explanation of perceived oscillation in a rotating trapezoid based entirely on dynamic properties of the projected image has been advanced by Guastella (1966). The observer is assumed to be most affected by the most rapidly changing element in the retinal image. The illusion is held to occur when the rate of change of the visual angle of the projection of the furthest edge is greater than that of the nearest edge, or when the change of rate of motion of the projection of the furthest edge is greater than that of the closest edge. Guastella did not provide empirical support for the specifics of this explanation, and recent results (Braunstein, 1971) indicate that the sources of information suggested by her are not the ones actually used by subjects in judging direction of rotation. Inequality of the vertical edges, which can cause the longer one to change more rapidly even when it is furthest from the observer, is not essential to the illusion. Horizontal changes appear to contribute to the illusion only when the vertical sides are not equidistant from the axis of rotation in the original form, a condition not met in Guastella's analysis.

While dynamic properties of optical stimulation are emphasized by Gibson (1966), he does not explain nonveridical perception of rotation on the basis of such properties. Gibson stated that his theory is "primarily a theory of correct perception [which] must explain incorrect perception by supplementary assumptions [1966, p. 287]." Correct perception is held to be achieved by observers through the use of invariants in the optic arrays. Perceptual processes based on invariants are analogous to algorithms in problem solving, as contrasted to the heuristic processes proposed here. As Gibson insists that motion perspective should produce veridical judgments of relationships in depth, he must resort to a postulate of misperceived static orientation to explain misperceived direction of rotary motion (Gibson & Gibson, 1957). This constitutes a major difficulty for Gibson's theory.

Johansson (1970), in a critique of Gibson's (1966) writings concerning "ecological optics," sketches a theoretical structure which stresses the analysis of the proximal stimulus by the visual system. This is contrasted to Gibson's emphasis on the relationship between the distal environment and the information available about this environment at the eye. Johansson stated that

> The efficiency of the [visual] system is given by a set of rules for stimulus data treatment (the

> programming of the visual computer, if you accept this metaphor), rules which work in an automatic way, but which result in veridicality when the proximal stimuli are projections from moving rigid objects and/or a rigid environment in motion relative to the eye. . . .
>
> Such rules have been shown experimentally to work in a blind, mechanical way and leave basically nothing for subjective choice. . . . We are dealing with both basic and rather rigid decoding principles (apparently common for all types of visual systems of retinal type) as one category and, as another, a more or less advanced ability to adapt the outcome from this primary decoding scheme when needed. This may perhaps be regarded as some kind of analogy to a basic programming together with a possibility of changing some instructions and constants within the program of a computer [pp. 73–74].

The present theory appears to fit into the principles outlined by Johansson. The only important deviation from the statements quoted above which has been proposed here is the insertion of "usually" before the words "results in veridicality." This would emphasize the belief expressed here that the rules used by the visual system are heuristic processes.

CONCLUSION

The study of the sensitivity of the human observer to transformations of the visual field has been characterized by an emphasis on finding stimulus characteristics which permit veridical perception of motion in depth (Gibson, 1957; Hay, 1966). It is essential, however, that theories of depth perception also explain nonveridical perception. As Leibowitz has pointed out, an unexplained illusion should "be considered an indication that our knowledge of fundamental laws of perception is inadequate [1965, p. 42]." The model presented here accounts for both veridical and illusory perception of rotating forms by postulating heuristic perceptual processes which, when applied in various stimulus situations, may yield either perceptual judgments which are objectively correct or are incorrect. It produces reasonable matches to results of experiments in which the actual or simulated viewing distance and the shape of the rotating form are varied and subjects are required either to give continuous direction judgments or to signal perceived oscillations. The present model is limited to rotating trapezoids and rectangles. This is a reflection of the availability of quantitative data for various stimulus configurations. Data are currently being collected on the perception of direction of rotation and figural coherence in three-dimensional dot patterns lacking visible contours. It is expected that the model can be generalized to include this case and other stimulus configurations.

REFERENCES

AMES, A. Visual perception and the rotating trapezoidal window. *Psychological Monographs*, 1951, **65**(7, Whole No. 324).

BRAUNSTEIN, M. L. Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology*, 1962, **64**, 415–420. (a)

BRAUNSTEIN, M. L. The perception of depth through motion. *Psychological Bulletin*, 1962, **59**, 422–433. (b)

BRAUNSTEIN, M. L. Sensitivity of the observer to transformations of the visual field. *Journal of Experimental Psychology*, 1966, **72**, 683–689.

BRAUNSTEIN, M. L. Motion and texture as sources of slant information. *Journal of Experimental Psychology*, 1968, **78**, 247–253.

BRAUNSTEIN, M. L. Perception of rotation in figures with rectangular and trapezoidal features. *Journal of Experimental Psychology*, 1971, **91**, 25–29.

BRAUNSTEIN, M. L., & PAYNE, J. W. Perspective and the rotating trapezoid. *Journal of the Optical Society of America*, 1968, **58**, 399–403.

COOK, T. H., MEFFERD, R. B., & WIELAND, B. A. Apparent reversals of orientation (perspective reversals) in depth as determinants of apparent reversals of rotary motion. *Perceptual and Motor Skills*, 1967, **24**, 691–702.

DAY, R. H., & POWER, R. P. Frequency of apparent reversal of rotary motion in depth as a function of shape and pattern. *Australian Journal of Psychology*, 1963, **15**, 162–174.

DAY, R. H., & POWER, R. P. Apparent reversal (oscillation) of rotary motion in depth: An investigation and a general theory. *Psychological Review*, 1965, **72**, 117–127.

GIBSON, J. J. Optical motions and transformations as stimuli for visual perception. *Psychological Review*, 1957, **64**, 288–295.

GIBSON, J. J. *The senses considered as perceptual systems.* Boston: Houghton Mifflin, 1966.

GIBSON, J. J., & GIBSON, E. J. Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology*, 1957, **54**, 129–138.

GRAHAM, C. H. On some aspects of real and apparent visual movement. *Journal of the Optical Society of America*, 1963, **53**, 1019–1025.

GRAHAM, C. H., & GILLAM, B. J. Occurrence of theoretically correct responses during rotation of the Ames Window. *Perception and Psychophysics*, 1970, **8**, 257–260.

GREEN, B. F., JR. Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 1961, **62**, 272–282.

GUASTELLA, M. J. New theory of apparent movement. *Journal of the Optical Society of America*, 1966, **56**, 960–966.

HAY, J. C. Optical motions and space perception: An extension of Gibson's analysis. *Psychological Review*, 1966, **73**, 550–565.

HERSHBERGER, W. A. Comment on "apparent reversal (oscillation) of rotary motion in depth." *Psychological Review*, 1967, **74**, 235–238.

JANSSON, G., & BORJESSON, E. Perceived direction of rotary motion. *Perception and Psychophysics*, 1969, **6**, 19–26.

JOHANSSON, G. Perception of motion and changing form. *Scandinavian Journal of Psychology*, 1964, **5**, 181–208.

JOHANSSON, G. On theories for visual space perception: a letter to Gibson. *Scandinavian Journal of Psychology*, 1970, **11**, 67–74.

LEIBOWITZ, H. W. *Visual perception*. New York: Macmillan, 1965.

MILES, W. R. Movement interpretation of a silhouette of a revolving fan. *American Journal of Psychology*, 1931, **43**, 392–405.

MURCH, G. M. The perception of rotary motion. *Journal of Experimental Psychology*, 1970, **86**, 83–85.

PASTORE, N. Some remarks on the Ames oscillatory effect. *Psychological Review*, 1952, **59**, 319–323.

POWER, R. P. Stimulus properties which reduce apparent reversal of rotating rectangular shapes. *Journal of Experimental Psychology*, 1967, **73**, 595–599.

WALLACH, H., & O'CONNELL, D. N. The kinetic depth effect. *Journal of Experimental Psychology*, 1953, **45**, 205–217.

WHITE, B. J., & MUESER, G. E. Accuracy in reconstructing the arrangement of elements generating kinetic depth displays. *Journal of Experimental Psychology*, 1960, **60**, 1–11.

ZEGERS, R. T. The reversal illusion of the Ames trapezoid. *Transactions of the New York Academy of Sciences*, 1964, **26**, 377–400.

(Received May 5, 1972)

Psychological Review
1972, Vol. 79, No. 6, 525–530

RELATION BETWEEN THE IMMEDIATE MEMORY SPAN AND THE MEMORY SEARCH RATE[1]

J. PATRICK CAVANAGH[2]

Université de Montréal

A comparison of memory span and item recognition data across several classes of stimuli shows that the time to search through a full memory (i.e., the number of items stored equals the memory span) in the item recognition task is a constant one-quarter of a second, regardless of material.

It has long been known that the span of immediate memory depends on the type of material tested (Jacobs, 1887); similarly, Sternberg (1969) has shown that the rate at which items are processed in memory in the item recognition task varies with the type of stimulus. To date, no comparison has been made between these two paradigms, perhaps because of their differences—the former is a recall task involving an error criterion and the latter a recognition task involving latency data.

Freund, Brelsford, and Atkinson (1969), however, have recently given evidence that differences in retrieval processes are sufficient to account for the differences between recall and recognition error performance. Furthermore, reaction time experiments comparing context recall and context recognition (Sternberg, 1969) indicate that the memory search processes in both tasks are identical. Thus, both error and latency measurements have supported the hypothesis that recall and recognition may use a common memory and store equivalent information. If this is the case, span and processing rate are both measures of the same memory system, and a comparison of the two across stimulus classes should point to important properties of the system.

Memory span has been measured by a great variety of methods (cf. Blankenship, 1938); generally, the task involves the presentation of increasingly longer lists of items, with the longest to achieve a certain error criterion in ordered recall being reported as the span. The span therefore can be considered a measure of the storage capacity of short-term memory for the class of material being tested.

The measurement of the memory processing rate has followed a more consistent procedure: subspan lists of varying length are memorized; a test item is then presented and classified as positive (a member of the memorized set) or negative (not a member). The dependent variable is the response latency from the presentation of the test item to the execution of the classification response; the independent variable is the length of the memorized list. In general, a linear relation has been found between these two variables (Sternberg, 1969). The slope of this relation thus estimates the processing rate (in milliseconds per item) for the class of stimuli being tested. No differentiation will be made between the two common variations in procedure (fixed vs. varied set) as both have been shown to give equivalent results and so presumably involve a common memory (i.e., short-term memory; Sternberg, 1969). Only a few of the studies reported here demonstrate significantly nonlinear relations (Egeth & Smith, 1967; Nickerson, 1966; Smith, 1967). In these cases, linear regressions estimate average processing rates, none of which differ markedly from others in the same stimulus class.

---

[1] This research was supported in part by Research Grant MH-07722 from the Public Health Service, National Institute of Mental Health.

[2] Requests for reprints should be addressed to J. Patrick Cavanagh, Département de psychologie, Université de Montréal, 90 Bellingham, Montréal 101, Québec.

I am grateful to William Chase, Sylvia Farnham-Diggory, and Stephen Checkosky for assistance in the preparation of this study and to Saul Sternberg for comments on an earlier version.

TABLE 1

PROCESSING RATE AND MEMORY SPAN AND CORRESPONDING STUDIES FOR VARIOUS STIMULUS CLASSES

Stimulus class	Processing rate[a]	Memory span[b]
Digits	35 (Bracey, 1969)	8.0 (Brener, 1940)
	27 (Burrows & Okada, 1971)	8.2 (Gates, 1916)
	30 (Cruse & Clifton, 1971)	7.3 (Kinsbourne & Cohen, 1971)
	38 (Sternberg, 1966)	7.7 (Olsson & Furth, 1966)
	39 (Sternberg, 1967)	7.3 (Watkins, 1914)[c]
	36 (Sternberg, 1969)	
	28 (Theios, Smith, Haviland, & Traupmann, 1971)	
	34 (Yio & Santa, 1970)	
M	33.4	7.70
Colors	38 (Checkosky, 1971)[d]	7.1 (Brener, 1940)
M	38.0	7.10
Letters	29 (Cavanagh & Chase, 1971)	7.3 (Brener, 1950)
	44 (Chase & Calfee, 1969)	5.4 (Kinsbourne & Cohen, 1971)
	53 (Chase & Posner, 1965)	
	43 (Cruse & Clifton, 1971)	
	41 (Ellis & Chase, 1971)	
	24 (Egeth & Smith, 1967)	
	42 (Forrin & Morin, 1969)	
	33 (Klatzky & Atkinson, 1971)	
	44 (Klatzky, Juola, & Atkinson, 1971)	
	65 (Nickerson, 1966)	
	26 (Williams, 1971)	
	38 (Wimberly, 1968)	
	39 (Yio & Santa, 1970)	
M	40.2	6.35
Words	52 (Burrows & Okada, 1972)	5.5 (Brener, 1940)
	50 (Goldring, 1968)	
	36 (Juola & Atkinson, 1971)	
	50 (Smith, 1967)	
M	47.0	5.50
Geometrical shapes	50 (Checkosky, 1970)	5.3 (Brener, 1940)
M	50.0	5.30
Random forms	92 (Bashark, 1968)	3.8 (Olsson & Furth, 1966)
	42 (Briggs & Blaha, 1969)	
	45 (Sternberg, 1969)	
	93 (Swanson, Johnsen, & Briggs 1972)	
M	68.0	3.80
Nonsense syllables	73 (Checkosky & Checkosky, 1972)	2.5 (Brener, 1940)
		4.3 (Watkins, 1914)[c]
M	73.0	3.40

[a] Milliseconds per item.
[b] Items.
[c] Based on 50% correct recall; values reported in study are based on 60% correct recall.
[d] Practice effect is reversed in color condition (shape filtering) of Checkosky's (1970) study; processing rates in Session 3 are greater than in Session 2. For this reason, Sessions 2 and 3 are averaged to obtain better reliability.

To compare the measures of the two tasks, a search of the memory span and memory search literature was conducted according to criteria that are described shortly. Seven classes of stimuli were found for which both measures were available: the digits 0 through 9; colors; the letters of the alphabet; familiar words; geometrical shapes such as

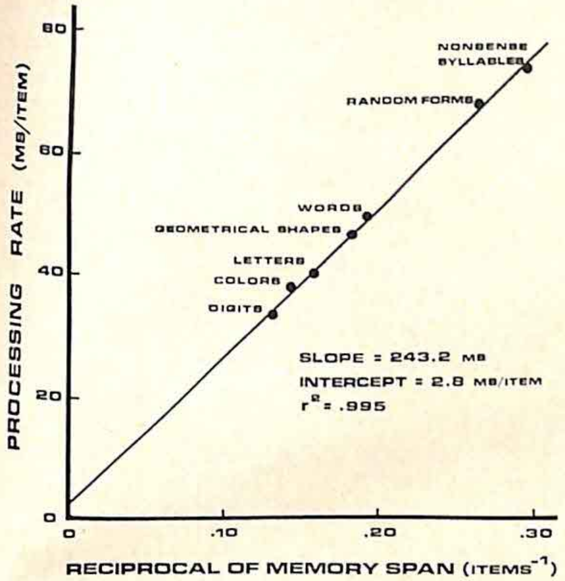


FIG. 1. The relation between short-term memory-processing rate and the reciprocal of the memory span for seven classes of stimuli (ms = millisecond).

squares, circles, and triangles; random forms—polygons of low associational value; and nonsense syllables (consonant-vowel-consonant).

The results are shown in Table 1. An inverse relationship is evident in the average values for the stimulus classes—the greater the memory span, the faster the processing rate. To determine the exact form of this relation, the mean values of processing rates are plotted, in Figure 1, against the reciprocals of the mean memory spans. The least squares linear fit accounts for 99.5% of the variance with a root mean standard deviation of 1.02 milliseconds/item ($df = 5$); the slope of the regression line is 243.2 milliseconds, and the intercept does not differ significantly from zero.

The data thus show a reciprocal relation between processing rate and memory span. The procedure for collecting the data attempted to rule out any confounding effects that might have arisen. In both paradigms, only results for adults were included. Since insufficient data on auditory item recognition are available, both memory span and item recognition studies were limited to visual modality. The items to be memorized could be presented either successively or simultaneously. The average exposure time per item in the studies selected was of the same magnitude (1–2 seconds) for both tasks. Reported memory span was the length of the longest list recalled correctly 50% of the time; although the method of scoring and the mode of recall—verbal, written, ordered selection from a large pool—varied between studies, the memory span values are not greatly affected by such variations (cf. Guilford & Dallenbach, 1925). Memory processing rates were estimated from the slope of the least squares linear relation between reaction time (averaged over positive and negative responses) and the length of the memorized list. Since the amount of practice given in memory span tasks is generally small, processing rate data were limited to the first two sessions of each study, where possible. Finally, since Klatzky, Juola, and Atkinson (1971) have shown that mixed presentation of test stimulus conditions results in processing that is markedly different from that found when a test stimulus type remains constant over a long series of trials, recognition studies were not accepted if any factors (modality, size, upper vs. lower case, etc.) other than list length or response type (positive or negative) varied within experimental blocks.

Although the criteria restrict the number of studies that can be drawn upon, the equating of experimental conditions and modality for the two tasks ensures an accurate determination of the memory span–processing rate relation. Since only group data are available in the literature, it is not possible at present to ascertain the degree to which the reciprocal relation holds for individual subjects. An appropriate within-subjects design might not only extend the results demonstrated here but might also indicate that the between-subjects variance in the two measures within stimulus classes is substantially accounted for by their reciprocal relation.

There are a number of theories that may predict this reciprocal relationship given suitable assumptions. The simplest of these might be termed the "size" hypothesis. According to this approach, the representation of a stimulus in memory is composed of a list of features. Short-term memory, with a constant, limited amount of available "space," can hold only a fixed number of features—T. The number of lists of features, that is, stimuli, of a particular class, c, that can be stored—$N(c)$—therefore depends on the number of features per list or the "size" of the representation—$s(c)$. The memory span, $N(c)$, is thus a measure of this size—the greater the mean number of features per stimulus, the fewer stimuli will fit in memory:

$$N(c) \cdot s(c) = T. \qquad [1]$$

Stimulus size is therefore inversely related to memory capacity as estimated by memory span:

$$s(c) = \frac{T}{N(c)}. \qquad [2]$$

If recognition involves serial testing of features (i.e., the test stimulus is compared feature by feature against each stored stimulus) and the time per feature test—t—is a constant, then the processing time per item for stimulus class c—$b(c)$—will be directly proportional to the number of features per item:

$$b(c) = s(c) \cdot t \qquad [3]$$

$$b(c) = \frac{T \cdot t}{N(c)}. \qquad [4]$$

The processing rate is thus inversely related to the memory span. Notice that an exhaustive search procedure similar to that described by Sternberg (1966) is necessary if this memory system is to be compatible with the data of the item recognition task.

A second hypothesis is based on the decay of memory traces. Mackworth (1963) has shown that the duration of report from a brief visual display is a constant 1 to 2 seconds, regardless of material. She concluded that this duration reflects the decay of the visual image and, therefore, that the number of items reported depends on the rate at which they can be identified. In a similar manner, this second hypothesis assumes that the duration of report from immediate memory when recalling a list of length equal to the memory span is also a constant—d—representing the decay time of items in short-term memory. The memory span for stimulus class c thus depends on the rate of report—$r(c)$—from immediate memory:

$$N(c) = \frac{d}{r(c)}. \qquad [5]$$

Since rehearsal is, in effect, self-report from immediate memory, the rate of rehearsal in the item recognition task will be assumed to be the same as the rate of report in the memory span task. The time to rehearse a list whose length equals the memory span is therefore also a constant, d, as given in Equation 5. If memory processing time for a test item is related by some function g to the time since its last rehearsal, then the average processing time when the memorized list length equals $N(c)$ is constant (the mean rehearsal-test delay will be one-half d, independent of stimulus class). Since processing time can be estimated (Sternberg, 1966) by the product of the processing rate, $b(c)$, and the number of items in memory, $N(c)$, the following expression results:

$$b(c) \cdot N(c) = g(d/2) \qquad [6]$$

$$b(c) = \frac{g(d/2)}{N(c)}. \qquad [7]$$

Since $g(d/2)$ is independent of c, a reciprocal relation between processing rate and memory span is again obtained. Note that (a) for reaction times to vary linearly with memory set size within stimulus classes, g itself must be a linear function; (b) additional assumptions are necessary to predict reaction times for negative test items—for example, negative reaction time could be proportional to the mean delay (averaged across memory items) between rehearsal of a memory item and presentation of the negative test; (c) memory span in this hypothesis is not viewed as storage capacity but only as the number of items that can be reported before the memory trace fades.

Finally, a third hypothesis, whose prediction of the inverse memory, span–processing rate relation led to this study, is based on a neural holographic model of short-term memory (Cavanagh, 1972). There is insufficient space here for a description of the model, but its prediction is based on the assumption that in dynamic activity, such as language or problem solving, the brain attempts to optimize its processing capacity. This implies, among other things, use of full memory capacity and minimization of recognition reaction times. The model allows reaction times in the item recognition task to be expressed as a function of a number of parameters involved in the representation and storage of information; minimized reaction times can therefore be determined by solving the total differentials of these equations when the number of items stored equals the memory span. The resulting minimized reaction times demonstrate an inverse memory span–processing rate relation.

These three hypotheses have involved assumptions that cannot be tested with the data surveyed, and they have been presented only to demonstrate possible explanations of the data. The reciprocal relation of memory span and processing rate does not, therefore, offer a critical test of the hypotheses but does provide a necessary boundary condition that must be met by any theory of the structure of memory. The corollary of this reciprocity condition is that the time required to process a full memory load is a constant, independent of the type of material stored, with a value in the order of one-quarter of a second, the slope of the regression line in Figure 1.

REFERENCES

Bashark, L. T. High-speed scanning of random figures in human memory. Unpublished master's thesis, Ohio State University, 1968.

Blankenship, A. B. Memory span: A review of the literature. *Psychological Bulletin*, 1938, **35**, 1–25.

Bracey, G. W. Two operations in character recognition: A partial replication. *Perception & Psychophysics*, 1969, **6**, 357–360.

Brener, R. An experimental investigation of memory span. *Journal of Experimental Psychology*, 1940, **26**, 467–482.

Briggs, G. E., & Blaha, J. Memory retrieval and central comparison times in information processing. *Journal of Experimental Psychology*, 1969, **79**, 395–402.

Burrows, D., & Okada, R. Serial position effects in high-speed memory search. *Perception & Psychophysics*, 1971, **10**, 305–308.

Burrows, D., & Okada, R. Parallel retrieval of physical and semantic information. Paper presented at the meeting of the Midwestern Psychological Association, Detroit, May 1972.

Cavanagh, J. P. Holographic processes realizable in the neural realm: Prediction of short term memory performance. Unpublished doctoral dissertation, Carnegie-Mellon University, 1972.

Cavanagh, J. P., & Chase, W. G. The equivalence of target and nontarget processing in visual search. *Perception & Psychophysics*, 1971, **9**, 493–495.

Chase, W. G., & Calfee, R. C. Modality and similarity effects in short-term recognition memory. *Journal of Experimental Psychology*, 1969, **81**, 510–514.

Chase, W. G., & Posner, M. I. The effects of auditory and visual confusability on visual and memory search tasks. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1965.

Checkosky, S. F. Speeded classification of multidimensional stimuli. Unpublished doctoral dissertation, State University of New York at Buffalo, 1970.

Checkosky, S. F., & Checkosky, N. B. Memory search for CVC and CCC trigrams. Unpublished manuscript, Lake Forest College, Lake Forest, Illinois, 1972.

Cruse, D., & Clifton, C., Jr. Recoding strategies and the retrieval of information from memory. (Cognitive Processes Laboratory Rep. No. 71–1) Amherst: University of Massachusetts, 1971.

Egeth, H., & Smith, E. E. On the nature of errors in a choice reaction task. *Psychonomic Science*, 1967, **8**, 345–346.

Ellis, S. H., & Chase, W. G. Parallel processing in item recognition. *Perception & Psychophysics*, 1971, **10**, 379–384.

Forrin, B., & Morin, R. E. Reaction times for items in short- and long-term memory. *Acta Psychologica*, 1969, **30**, 126–141.

Freund, R. D., Brelsford, J. W., Jr., & Atkinson, R. C. Recognition vs. recall: Storage or retrieval differences? *Quarterly Journal of Experimental Psychology*, 1969, **21**, 214–224.

Gates, A. I. The mnemonic span for visual and auditory digits. *Journal of Experimental Psychology*, 1916, **1**, 393–403.

Goldring, S. L. The effect of presentation and test mode on short term retention of words. Unpublished master's thesis, University of Wisconsin, Madison, Wisconsin, 1968.

Guilford, J. P., & Dallenbach, K. M. The determination of memory span by the method of constant stimuli. *American Journal of Psychology*, 1925, **36**, 621–628.

Jacobs, J. Experiments on "prehension." *Mind*, 1887, **12**, 75–79.

JUOLA, J. F., & ATKINSON, R. C. Memory scanning for words versus categories. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10**, 522–527.

KINSBOURNE, M., & COHEN, V. English and Hebrew consonant memory span related to the structure of the written language. *Acta Psychologica,* 1971, **35**, 347–351.

KLATZKY, R. L., & ATKINSON, R. C. Specialization of the cerebral hemispheres in scanning for information in short-term memory. *Perception & Psychophysics,* 1971, **10**, 335–338.

KLATZKY, R. L., JUOLA, J. F., & ATKINSON, R. C. Test stimulus representation and experimental context effects in memory scanning. *Journal of Experimental Psychology,* 1971, **87**, 281–288.

MACWORTH, J. F. The duration of the visual image. *Canadian Journal of Psychology,* 1963, **17**, 62–81.

NICKERSON, R. S. Response times with a memory-dependent decision task. *Journal of Experimental Psychology,* 1966, **72**, 761–769.

OLSSON, J. E., & FURTH, H. G. Visual memory span in the deaf. *American Journal of Psychology,* 1966, **79**, 480–484.

SMITH, E. E. Effects of familiarity on stimulus recognition and categorization. *Journal of Experimental Psychology,* 1967, **74**, 324–332.

STERNBERG, S. High-speed scanning in human memory. *Science,* 1966, **153**, 652–654.

STERNBERG, S. Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics,* 1967, **2**, 45–53.

STERNBERG, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist,* 1969, **57**, 421–457.

SWANSON, J. M., JOHNSEN, A. M., & BRIGGS, G. E. Recording in a memory search task. *Journal of Experimental Psychology,* 1972, **93**, 1–9.

THEIOS, J., SMITH, P. G., HAVILAND, S., & TRAUPMANN, J. Memory scanning is a serial self-terminating process. (Wisconsin Mathematical Psychology Program Rep. No. 71–1) Madison, Wis.: University of Wisconsin, 1971.

WATKINS, S. H. Immediate memory and its evaluation. *British Journal of Psychology,* 1914, **7**, 319–348.

WILLIAMS, J. D. Memory ensemble selection in human information processing. *Journal of Experimental Psychology,* 1971, **88**, 231–238.

WIMBERLY, F. C. A generalization of some results of Sternberg. (Complex Information Processing Working Paper No. 135) Pittsburgh, Pa.: Carnegie-Mellon University, 1968.

YIO, J. H., & SANTA, J. L. Reaction time in short-term recognition with digits and letters. *Psychonomic Science,* 1970, **20**, 121–122.

(Received May 25, 1972)

Psychological Review
1972, Vol. 79, No. 6, 531–535

THEORETICAL NOTES

CONFIDENCE RATINGS IN RECALL: A REANALYSIS[1]

DOUGLAS L. HINTZMAN[2]

University of Oregon

Confidence rating data from two paired-associate learning experiments are reanalyzed, and two arguments are made. The primary argument is that Bernbach's claim that Type 2 d' does not increase with practice is false, and this can be shown if pairs with different past histories are analyzed separately. The secondary argument is that when pooled group data are used to construct Type 2 operating characteristics, as has been the practice, subject selection tends to confound effects of response bias with those of discriminability. Data from the two experiments are analyzed using both the group Type 2 analysis and an alternative analysis that is free from subject selection. Outcomes of both analyses show that the more often responses have been given correctly in the past, the more easily subjects discriminate them from errors.

A current skirmish in the incremental versus all-or-none learning controversy involves the use of Type 2 operating characteristics to analyze confidence rating data in paired-associate recall. The Type 2 operating characteristic differs from Type 1 in that "signal" and "noise" trials are defined by the performance of the subject himself, rather than by a manipulation under direct control of the experimenter. Thus in a recall task the subject first gives a response and then rates his confidence in the correctness of that response. The Type 2 operating characteristic plots P (confidence $\geq Y|$correct) against P (confidence $\geq Y|$error), where Y is a numerical rating of confidence that the response is correct. The distance of this curve from the main diagonal, d', reflects the subject's ability to discriminate his own correct responses from his errors.

Bernbach (1967) has proposed that if values of P (confidence $\geq Y|$correct) are adjusted for guessing before operating characteristics are plotted, the resulting d' can be interpreted as reflecting the subject's ability to discriminate between learned and unlearned items. Data (Bernbach, 1967; Bernbach & Bower, 1970) seem to indicate that the discriminability of learned and unlearned items does not increase with practice. Such an outcome is what would be expected if learning were a matter of all-or-none transfer of a pair from an "unlearned" to a "learned" state. The finding has therefore been taken as evidence for a finite state, and against a continuous strength representation of the memory trace (Bernbach, 1967, 1971).

The primary argument to be made here is that contrary to previous conclusions, such discriminability does increase with practice, and this can be shown if items with different past histories are considered separately. A secondary argument is that when Type 2 operating characteristics are constructed from pooled group data, as has been the practice, subject selection tends to confound effects of response bias with effects of discriminability. This defeats the purpose of signal detection analysis, and makes interpretation difficult. Because of the problem, data are analyzed here first using a group Type 2 analysis similar to Bernbach's, and second using a procedure that is not susceptible to subject selection artifacts. Both analyses lead to a rejection of Bernbach's two-state memory model.

Two sets of paired-associate data with confidence ratings were selected for reanalysis. They are referred to here as Experiments I and II. Data of Experiment I, an unpublished study of my own, were used by Bernbach

---

[1] The research reported herein was performed pursuant to a grant with the Office of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

[2] Requests for reprints should be sent to Douglas L. Hintzman, Department of Psychology, University of Oregon, Eugene, Oregon 97403.

TABLE 1

TYPE 2 RECEIVER OPERATING CHARACTERISTIC ANALYSIS, GROUP d_s VALUES

Trial	Correct-response Category				All Correct	
	First	Second	Third	Fourth	Unadjusted	Adjusted
Experiment I						
1	1.55				1.55	3.05
2	1.50	3.20			1.85	2.20
3	1.20	2.75	4+		2.10	2.30
4	1.10	1.55	2.30	3.30	1.75	1.80
Experiment II						
1	.70				.70	1.05
2	.85	2.30			1.25	1.60
3	1.10	1.60	3.10		1.50	1.60
4	.65	1.25	1.95	3.05	1.45	1.55

(1967) to support his original argument. In that study, each subject learned a 16-pair list using the study-test procedure. After giving a response to each stimulus item, the subject rated his confidence in the correctness of his answer on a scale from 0 to 10. There were three conditions in all, involving lists with 4, 8, or 16 different response alternatives. The 16-alternative data ($N = 15$) were selected to be reanalyzed here because according to Bernbach's model (1967, p. 477) the probability of a correct guess in the unlearned state should be fairly low; $g = .063$. The data of Experiment II are from a published study in which each subject learned a 14-pair list using the anticipation procedure, with unpaced recall intervals (Hintzman, 1967). Confidence ratings, again, ranged from 0 to 10. The complete study orthogonally manipulated stimulus similarity (high vs. low) and number of response alternatives (2 vs. 14). The low-14 data ($N = 16$), with an a priori guessing rate of $g = .071$, were chosen for reanalysis.

For both experiments, Type 2 operating characteristics were constructed from the pooled group data. A d_s value (an estimate of d') was found for each curve by doubling the absolute value of the z score for false alarms at the point where the curve crosses the negative diagonal (Banks, 1970, p. 85). Two types of curves were constructed in which all correct responses were categorized together: one was an ordinary Type 2 analysis, while the other used Bernbach's correction for guessing. The resulting discriminability measures are presented in the two right-hand columns of Table 1. The adjusted values estimate the discriminability of the two memory states of Bernbach's model. As Bernbach has reported, overall discriminability did not increase with trials. (The adjusted values for Experiment I agree with the graph of the same data presented by Bernbach [1967, Figure 14].)

A different picture emerged, however, when correct responses were further categorized according to their past histories. Separate operating characterisics were plotted for items gotten correct different numbers of times. The differences among the resulting curves were quite apparent; within trials, discriminability increased monotonically with he number of previous correct responses. This analysis, like the combined analyses, is summarized in terms of d_s values in Table 1. The d_s values for these separate categories were not adjusted for guessing using Bernbach's formula, since the probability that a correct response was a guess is not independent of the past history of the pair. In fact, one might wish to argue for this reason that a within-trials increase in d_s is not inconsistent with the two-state model—the more times an item has been gotten correct, the less likely it is that it is in the unlearned state, and therefore the higher the average confidence. But this explanation will not work, because only the first correct response category could contain an appreciable proportion of guesses, while the increases across second, third, and fourth correct responses were both large and consistent. Note especially that in both experiments the discriminability of third and fourth correct responses from errors was considerably greater than the discriminability of the learned and unlearned states according to Bernbach's model (last column). The particular pattern of the within-trials increase in Type 2 discriminability shown in Table 1 appears to be inconsistent with the two-state memory model.

Apparently, the increase in Type 2 discriminability with practice was not noticed before because in previous analyses all correct responses were combined, and comparisons were made between, rather than within, trials. Other things equal, the increase would be expected to show up in between-trials comparisons, but it was balanced by a tendency for discriminability to decrease over trials within a given correct-response category (i.e., reading down columns of Table 1). Why did this occur? At least two possible causes can be suggested. Later-learned items, being more difficult, may enter the correct-response pool at lower discriminability than items learned on early trials. Alternatively, fast learners may be able to discriminate correct responses from errors better than slower learners. Since, in the group Type 2 analysis, data are pooled over subjects, both item selection and subject selection may confound the comparison of performance on one trial with that on another. Under the circumstances, any effect or lack of effect of trials on discriminability could be artifactual and should be interpreted with utmost caution.

Whatever problems may be involved in such between-trial comparisons, a more subtle selection problem can be shown to exist in the within-trial comparisons upon which the group Type 2 discriminability measure itself is based. Item selection is not a problem here, since it is item selection—the selective comparison of correct responses and errors—that is the basis of the Type 2 analysis. But subject selection is another matter. Each subject-item contributes one rating to the data pool, and this rating is categorized according to whether the response was correct or incorrect (in the present case correct responses were further categorized according to their past histories). Different subjects, according to their abilities, contribute differentially to these categories. Thus the item selection, which is deliberate, may be confounded by subject selection. The seriousness of this problem arises from the fact that the derivation of *d′* assumes that confidence ratings are based on the same criteria, independently of whether the response was correct or incorrect. To the extent that correct responses and errors are made by different subjects, there is no guarantee that this is the case.

To illustrate, consider an artificial example in which data generated by five subjects are submitted to a group Type 2 analysis like that summarized in Table 1. Assume that the subjects differ in learning ability and that a given subject learns all pairs on the same trial. This occurs in such a way that on Trial 4 the responses are distributed as follows: Subject 1 makes all the fourth correct responses, Subject 2 the third, Subject 3 the second, and Subject 4 the first. Subject 5, the poorest learner of the group, makes all the errors against which the other subjects' correct responses are plotted. Even if none of the individual subjects is able to distinguish correct responses from errors, differences among the subjects in use of the confidence scale can produce nonzero "*d′*" values. For example, if subjects' confidence were positively correlated with learning ability, then obtained *d′* values would be positive and would increase with number of times correct (as in Table 1). If the correlation between confidence and ability were negative, a decreasing function, with negative *d′* values, would result. Various irregular relationships would produce still other outcomes. Furthermore, whatever the nature of the subject differences, their effects would be exaggerated by the fact that each subject contributes several responses per trial.

The point is that group Type 2 curves do not necessarily reflect the performance of individual subjects. In the kind of pooled group analysis used by Bernbach (1967), and used here in Table 1, the effects of discriminability are confounded with effects of response bias. To the extent that subject selection is present, one cannot be certain what combination of discriminability and bias produced a given result. (Differences among subjects' criterion setting could not only artifactually raise or lower discriminability measures but might also be responsible for some otherwise puzzling findings regarding operating characteristic slopes—e.g., Donaldson & Glathe, 1970).

The problem is of concern here because the degree of subject selection was rather substantial in the present experiments. Over the first four trials of Experiment I, out of a total of 64 responses per subject, the number of errors ranged from 8 for one subject to 50 for another. In Experiment II, with 56 responses per subject, the range was from 13 to 54. Considering the possible effects of this variability, not even

TABLE 2

CONFIDENCE RATINGS: MEANS OF SUBJECT MEANS

Trial	Error	Correct response			
		First	Second	Third	Fourth
Experiment I					
1	.91 (15)	4.88 (15)			
2	1.37 (15)	5.97 (15)	9.50 (11)		
3	1.74 (14)	5.14 (14)	8.67 (15)	9.91 (11)	
4	2.90 (14)	6.80 (12)	8.05 (14)	8.95 (15)	9.80 (10)
Experiment II					
1	1.60 (16)	3.61 (15)			
2	2.54 (16)	4.06 (14)	8.35 (11)		
3	3.08 (15)	6.31 (15)	7.61 (10)	9.87 (10)	
4	3.61 (15)	5.32 (9)	7.08 (15)	8.67 (11)	9.95 (10)

Note.—Number of subjects contributing shown in parentheses. For Experiment I, $N = 15$; for Experiment II, $N = 16$.

the large and consistent within-trials increase in d_s shown in Table 1 can be safely assumed not to be an artifact of the treatment of group data as though they were from a single subject. Thus while the group Type 2 analysis is useful because it allows a comparison with previous data, it does not prove the two-state model wrong. One could still argue that Bernbach's theoretical position regarding invariance of Type 2 discriminability is basically correct, and reject the data in Table 1 as contaminated by subject selection artifacts. It seems worthwhile, therefore, to seek an alternative method of analysis that is not subject to the same criticism.

The ideal way to overcome the subject selection problem would be to do separate Type 2 analyses for each subject, and then examine average d_s values. However, this procedure requires a very large number of observations per subject per trial, since confidence ratings on correct responses must partly overlap those on errors in order for an operating characteristic to be plotted. In neither of the present experiments are individual Type 2 analyses possible.

An alternative solution, which should serve the present purpose of verifying or disconfirming the within-trials increase in d' with number of previous correct responses, is based on confidence rating means. If on a given trial an individual subject's mean confidence on correct items increases regularly with number of times correct, then his ability to distinguish those correct responses from errors must be increasing also. Accordingly, the data were again reanalyzed, this time computing for each subject, where possible, the mean confidence ratings on errors, and on first, second, third, and fourth correct responses. The means of the subject means for each category on each trial are shown for both experiments in Table 2. Note that the means presented in Table 2 are not entirely free from the subject selection of Table 1; but here a subject could contribute no more than one value to each category (the number of subjects contributing to each is shown in parentheses). Furthermore, statistical analyses of individual subjects' data indicate that the pattern displayed by the means of Table 2 was in fact representative of performance of single subjects. For each contributing subject, on each of Trials 2–4, the difference in mean confidence between the ith and the jth correct response categories was computed, where i was the highest category for the subject on that trial and j the lowest, $i \neq j$. For example, if a given subject on Trial 4 made second, third, and fourth, but no first correct responses, the subject's score was the difference in mean confidence between the fourth and second correct response categories. Tests of the hypothesis that the mean difference was zero yielded the following results: For Experiment I, Trials 2, 3, and 4, respectively: $t(10) = 5.49$; $t(14) = 4.50$, both $ps < .001$; $t(14) = 4.04$, $p < .01$. For Experiment II: $t(10) = 3.88$; $t(11) = 3.15$, both $ps < .01$; $t(13) = 2.64$, $p < .05$.

Thus the conclusion that was tentatively drawn from Table 1, but that was suspect for the reasons discussed previously, is confirmed. According to this measure, which is free of subject selection, the ability to distinguish correct responses from errors on a given trial increased directly with the number of previous correct responses on the pair. Again, the magnitude of the increase and its consistency cannot be explained in terms of different proportions of guesses in the different categories, since only first-correct responses could include an appreciable proportion of guesses. Further, mean ratings to the best-learned items were extremely close to the maximum of 10; and that would not be the case if those items were in an undifferentiated learned state from which ratings are distributed widely over the scale.

What bearing does the present reanalysis have on the continuous strength versus finite state question? One conclusion that appears inescapable is that the outcome is inconsistent with the simple two-state model for confidence ratings proposed by Bernbach (1967). Beyond this, however, there are few definite conclusions that can be made. The confidence rating data appear, at least qualitatively, to be consistent with a large number of alternative hypotheses. Two of them seem to be worth mentioning: First, Bernbach (1971) considered and rejected two versions of a continuous strength model—one of them based confidence on the strength of the retrieved association and the other on the difference in strengths between the retrieved and the next strongest associations. The reason for rejecting the model (invariance of Type 2 discriminability) now appears to be invalid. Second, the present result seems to be more consistent, in spirit at least, with Bernbach's (1970a, 1970b) "replica" model than was his original conclusion. The replica model assumes that after an item is first learned additional copies or replicas of its trace may be stored due to rehearsal or repetition. It seems natural to assume that confidence is a function of the number of replicas of an item in storage, and that would explain the present result. A similar notion is that confidence is based in part on a judgment of the frequency with which the retrieved response was confirmed on past trials as correct. Such frequency information could, logically, be stored independently of the memory state underlying retrieval.

REFERENCES

BANKS, W. P. Signal detection theory and human memory. *Psychological Bulletin,* 1970, **74**, 81–99.

BERNBACH, H. A. Decision processes in memory. *Psychological Review,* 1967, **74**, 462–480.

BERNBACH, H. A. Replication processes in human memory and learning. In J. T. Spence & G. H. Bower (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 3. New York: Academic Press, 1970. (a)

BERNBACH, H. A. A multiple-copy model for post-perceptual memory. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970. (b)

BERNBACH, H. A. Strength theory and confidence ratings in recall. *Psychological Review,* 1971, **78**, 338–340.

BERNBACH, H. A., & BOWER, G. H. Confidence ratings in continuous paired-associate learning. *Psychonomic Science,* 1970, **21**, 252–253.

DONALDSON, W., & GLATHE, H. Signal-detection analysis of recall and recognition memory. *Canadian Journal of Psychology,* 1970, **24**, 42–56.

HINTZMAN, D. L. Some tests of a discrimination net theory: Paired-associate learning as a function of stimulus similarity and number of responses. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6**, 809–816.

(Received December 31, 1971)

Psychological Review
1972, Vol. 79, No. 6, 536–537

CONFIDENCE RATINGS FOR INDIVIDUAL ITEMS IN RECALL

HARLEY A. BERNBACH[1]

Purdue University

Hintzman claims, as a result of a 1972 reanalysis of confidence rating data in paired-associate learning, that Bernbach's 1967 claim that d^* is independent of trial number is false. His conclusion is simply wrong; he has rediscovered the finding that item differences contribute to the variability of apparent oldness (or strength, familiarity, or the like) for remembered and not-remembered items. He has, in fact, provided additional evidence for the well-known invariance of d^* with changes in memory variables such as practice or serial position.

Subjects generally generate a range of confidence ratings when asked to rate confidence in both recall and recognition tasks. This appears to be because the correct/incorrect or old/new sets of items have confidence distributions on the decision scale (strength, apparent oldness, or the like) that have nonzero variances. It is this fact that makes a signal detection analysis of such data both possible and necessary. The source of this variance has generally been referred to as noise or decision variability, but it would be a mistake to assume that it is merely error variance in a homogeneous set of items. It is just this mistake that has led Hintzman to a totally incorrect conclusion from his reanalysis of some confidence rating data from recall experiments.

The fact is that individual items differ in their mean locations on the recall confidence scale both when they are remembered and when they are not. The evidence for this is hardly new. It has been known for a long time that individual items in recognition memory differ in their locations on the decision scale in similar fashion in the distributions for both new and old items. Consider the continuous short-term recognition memory task first suggested by Shepard and Teghtsoonian (1961). In this task, each item appears twice in a continuous string of items and thus is tested both when it is new and when it is old. It has been noted (e.g., Melton, Bernbach, & Reicher, 1964; Melton, Sameroff, & Shubot, 1963) that an item is far more likely to be called old on its second appearance if it was called old (incorrectly) on its first appearance than if it was correctly identified as new when it was first presented. A formal theoretical statement that different item locations on the decision scale are the basis for this phenomenon was presented by Bernbach (1964). This issue was not pursued, primarily because it did not discriminate among theories, as it is consistent with either a strength or finite state theory.

It is an easy theoretical step to add the assumption that items high on the decision scale are learned fastest. This assumption permits a complete explanation of Hintzman's (1972) data without discriminating between continuous and finite state theories. That is, if this assumption holds, the set of items that has been correct several times in a row will have a mean scale location, relative to the set of errors, that is higher than that for the set of items correct for the first time. It is the mean relative location that Hintzman calls "discriminability" and measures with the index, d_s. Both types of theory would predict a decrease in this measure as one reads down columns in Hintzman's Table 1. In both cases, what is being measured is not the discriminability of these subsets of correct items but their mean scale location relative to *all* the incorrect items. What is important for purposes of verifying the finite state theory is that a proper index of the discriminability of all correct from incorrect items must remain constant across trials. This appears to be the case both in Hintzman's Experiment I, as had been noted previously (Bernbach, 1967), and Experiment II. The result does not appear to be compatible with strength theory (Bernbach, 1971a), despite Hintzman's appeal to an accidental "balance" between the increase in his d_s values across rows and the unexplained decrease reading down columns. This balance could hardly be accidental, considering the considerable number of replications of the variance of d^* (the mean discriminability for all items in this instance) that have been reported (cf. Bernbach, 1971b). In fact, the values of d_s

---

[1] Requests for reprints should be sent to Harley A. Bernbach, Department of Psychological Sciences, Purdue University, Lafayette, Indiana 47907.

from Experiment II, appropriately corrected for guessing, actually lend support to the claim that d^* is invariant with moderate amounts of practice.

Hintzman's second argument, backed by the data in Table 2, needs only brief consideration. Bernbach and Bower (1970) demonstrated clearly that mean confidence rating can change with d' constant, as a function of criterion change only. They also found that cues for criterion shifts can be picked up by subjects within a list. There are many possible reasons for subjects to have used different criteria for the different items with which Hintzman is concerned; Hintzman suggests at least one of these in his own discussion. Whatever the case, mean confidence rating is a worthless summary statistic in any instance except one in which it is absolutely certain that there is no taint of criterion change, a rare situation indeed.

Finally, since Hintzman has raised the point of possible problems with using data pooled across subjects, it should be pointed out that it would be unlikely that we find d^* invariant so repeatedly if the suggested artifact were working. Hintzman's argument is that the pooling of data might be the cause of a change in d^* with practice, a change that does not occur. If Hintzman is really concerned about the effects of pooling data, it would seem that he should collect data from individual subjects rather than perform a faulty reanalysis of some old (pooled) data.

REFERENCES

BERNBACH, H. A. A decision and forgetting model for recognition memory. (Tech. Rep. No. 64-4) Ann Arbor: Michigan Mathematical Psychology Program, 1964.

BERNBACH, H. A. Decision processes in memory. *Psychological Review,* 1967, **74**, 462–480.

BERNBACH, H. A. Strength theory and confidence ratings in recall. *Psychological Review,* 1971, **78**, 338–340. (a)

BERNBACH, H. A. The invariance of d^* in memory. *Psychological Bulletin,* 1971, **76**, 149–150. (b)

BERNBACH, H. A., & BOWER, G. H. Confidence ratings in continuous paired-associate learning. *Psychonomic Science,* 1970, **21**, 252–253.

HINTZMAN, D. L. Confidence ratings in recall: A reanalysis. *Psychological Review,* 1972, **79**, 531–535.

MELTON, A. W., BERNBACH, H. A., & REICHER, G. M. Stimulus frequency in continuous short-term recognition memory. Paper presented at the meeting of the Psychonomic Society, Niagara Falls, Ontario, October 1964.

MELTON, A. W., SAMEROFF, A., & SHUBOT, E. D. Short-term recognition memory. Paper presented at the meeting of the Psychonomic Society, Bryn Mawr College, August 31, 1963.

SHEPARD, R. N., & TEGHTSOONIAN, M. Retention of information under conditions approaching a steady-state. *Journal of Experimental Psychology,* 1961, **62**, 302–309.

(Received August 16, 1972)

Psychological Review
1972, Vol. 79, No. 6, 538–539

GREGORY'S THEORY OF ILLUSIONS:

SOME DISCONFIRMING EVIDENCE IN THE CASE OF THE MÜLLER-LYER ILLUSION[1]

LESTER E. KRUEGER[2]

City College of the City University of New York

According to R. L. Gregory, the portion of the Müller-Lyer figure having the outgoing fins is seen in perspective as representing an inside corner, and thus, through misplaced size constancy, the line from which the fins extend outward is perceived as longer. Two findings argue against the perspective notion as a complete explanation of the illusion. First, as the outgoing fins are lengthened, the illusion at first increases in size, until the fins reach a certain length, beyond which the illusion begins to decrease in size as the fins are further lengthened. Second, the same illusion, and the same nonmonotonic effect of fin length, is found when, at each end of the line, the two fins are collapsed together to form simply a demarcated appendage to the line. An explanation based on figure-ground separation is put forward. Short fins are not perceptually distinct from, but rather are grouped with, the line from which they extend, so that lengthening these fins causes an increase in the perceived extent of the line. Lengthening the fins beyond a certain point, however, causes an ever clearer figure-ground separation of the line from the fins, thus negating more and more of the effect of the fins.

Gregory (e.g., 1966) has invoked the notion of misplaced constancy to explain the Müller-Lyer and other visual illusions. In the portion of the Müller-Lyer figure in which the fins are outgoing (i.e., extend away from the line to which they are attached), the fins, he notes, could represent the inside corner of a room, and the resulting depth cue, even though countermanded by other cues, might feed into the size-constancy mechanism and thus increase the perceived length of the inside-corner line.

Two findings, however, seem to point to a more peripheral and less cognitive explanation of the Müller-Lyer illusion. First, when the part of the Müller-Lyer figure with the outgoing fins is tested, a particular fin length produces a maximal effect (Heymans, 1895–1896; Lewis, 1909; Obonai, 1954). Keeping the angle between fins constant, and keeping it below 50°, as the fins are successively lengthened, the center horizontal line located between the two pairs of fins increases in apparent length, but only until fin length reaches a certain value, after which further increases in fin length decrease the apparent length of the center line. The maximal effect seems to occur when the fin length is such that its normal projection onto the hypothetical extension of the center horizontal line (that is, the portion of the imaginary extension that is cut off when a perpendicular line is dropped to it from the end of a fin) is half the length of the center line.

The second finding, emphasized by Lewis (1909) and Obonai (1954), is that a similar illusion is obtained with a figure in which the angle between the outgoing fins is reduced to zero (i.e., each pair of fins is collapsed together to form a single line) and a small vertical mark is inserted to indicate where the center line ends and the fins begin: the center line appears longer and longer as the collapsed fins are lengthened, but only up to a certain point, after which further increase in length decreases the illusion.

Recently, this investigator attempted to draw a closer connection between the illusions arising from these two types of figures by showing that practice or fatigue has the same effect in both cases. On each trial, the subject had to adjust the length of one horizontal line to equal the length of a horizontal line which was located to its left and at whose ends were outgoing fins at an angle of incidence of either 30° or 0° (i.e., collapsed fins). For the figure with the collapsed fins, instead of vertical marks there was a short spatial separation between the center line and the fin on each side. A clear effect of practice was not found for either figure over three one-hour sessions with six subjects, but for both types of figure a maximal illusion was found when the projected length (30° case) or the actual length (0° case) of the outgoing fins was about half that of the center line.

---

[1] This study was supported by NIH Grant HD04869-02.

[2] Requests for reprints should be sent to Lester E. Krueger, Department of Psychology, City College of the City University of New York, New York, New York 10031.

The two findings seem difficult to account for in terms of misplaced constancy. The depth cue produced by the outgoing fins ought to become ever larger as the fins are lengthened, rather than to increase at first but then to become weaker as fin length passes a particular value. Further, if, as the evidence suggests, the same illusion is produced by a figure containing horizontal, collapsed fins, as by a figure with regular diagonal fins, then any argument based on perspective or size constancy seems even less appropriate as a complete explanation of the illusion.

Several more peripheral and less cognitive explanations might account for the two findings. Lewis (1909) and Obonai (1954) rely on assimilation and contrast; short outgoing fins are assimilated to the center horizontal line, making it appear longer, whereas long outgoing fins are contrasted with the center horizontal line, making it appear shorter. An explanation based on figure-ground separation also is possible. As Beck (1966) and Krueger (1970) have shown, a letter is more difficult to identify when irrelevant visual material barely surrounds it than when it lies within a larger surrounding irrelevant context, perhaps because, as Beck suggests, adding more irrelevant material leads to greater figure-ground separation between the letter (figure) and the irrelevant material (ground). In the Müller-Lyer illusion, when the outgoing fins are short, the outgoing fins and the portion of horizontal line between them may cohere as a single figure, with the fin length being assimilated to that of the horizontal line, whereas when the outgoing fins are lengthened still further, figure-ground separation may tend to occur, so that the fins become more and more distinct as separate entities and have less and less effect on the horizontal line.

REFERENCES

BECK, J. Effect of surround size on the perception of texture patterns. *Journal of Experimental Psychology*, 1966, 72, 68–75.

GREGORY, R. L. *Eye and brain.* New York: McGraw-Hill, 1966.

HEYMANS, G. Quantitative Untersuchungen über das "optische Paradoxen." *Zeitschrift für Psychologie*, 1895–1896, 9, 221–255.

KRUEGER, L. E. Effect of bracketing lines on speed of "same"–"different" judgment of two adjacent letters. *Journal of Experimental Psychology*, 1970, 84, 324–330.

LEWIS, E. O. Confluxion and contrast effects in the Müller-Lyer illusion. *British Journal of Psychology*, 1909, 3, 21–41.

OBONAI, T. Induction effects in estimates of extent. *Journal of Experimental Psychology*, 1954, 47, 57–60.

(Received January 6, 1972)

Psychological Review
1972, Vol. 79, No. 6, 540–541

COMMENTS ON L. E. KRUEGER'S "DISCONFIRMING EVIDENCE" OF R. L. GREGORY'S THEORY OF ILLUSIONS

RICHARD L. GREGORY

University of Bristol

The author responds to L. E. Krueger's 1972 suggestion that a more peripheral and less cognitive type of explanation is indicated by experimental facts for visual distortions such as those given by the Muller-Lyer figure. In 1963, R. L. Gregory had suggested a theory of "constancy scaling" based on a missetting of the size-scaling mechanism. It is argued that perceptual scale distortions are the problem being dealt with, and that in principle a mechanism capable of producing size scale changes is required. Constancy scaling is defended as a logical candidate.

Krueger (1972) raised several points concerning versions of the Muller-Lyer illusion which seem relevant to the "inappropriate constancy scaling" theory of distortion illusions. The theory (Gregory, 1963) is that distortions, such as given by the Muller-Lyer figure, are due to missetting of the size-scaling mechanism which normally maintains apparent size constant over a wide range of object distance, by effectively compensating shrinking of the retinal image with increasing object distance. It is supposed that perspective features of the retinal image set apparent size, converging lines giving expansion. In the illusion figures (and all perspective pictures), typical depth-related features, such as perspective convergence, are present on a flat surface. Since the distortions occur though the figures are seen as flat, it is supposed that size scaling can be set *directly* by these features, quite apart from depth as perceived at that time. This is called "depth cue scaling."

This is not the only kind of scaling (or way in which scaling can be set or misset), for size changes occur with depth reversals of ambiguous objects, especially when there is no visible background to prevent the figure being seen as three dimensional. Since these size changes occur with changes in perceived depth (increase in size with increasing apparent distance), though the retinal image remains unchanged, this scaling must be set *not* by retinal features but by the prevailing hypothesis, or perception, of depth. Therefore, it must be central, and cognitive, as it is tied to prevailing object perception, not to the prevailing sensory stimulus pattern. This is called "depth hypothesis scaling."

Depth cue and depth hypothesis scaling are essentially different, but an error in either must, unless somehow compensated, produce a corresponding error in perceived size. Converging lines lying on a plane give illusory expansion where they (or enclosed features) would normally be more distant; and with depth reversals, size increases with apparent distance, as in wire cubes or luminous Necker cubes.

Krueger is worried that some experimental facts about the Muller-Lyer illusion—especially that there is an optimum length of fin rather than monotonic increase in distortion with fin length—suggest a more peripheral and less cognitive type of explanation. However, the cue scaling, relevant here, is not conceived to be at a high cognitive level, as it is set directly by (perspective) features of the figure, apart from perceived depth or identification of the pattern as an object. This puts cue scaling at roughly the logical level of ethological "releasers": they would hardly be said to be cognitive. To say that the "depth cue produced by the outgoing fins ought to become ever larger as the fins are lengthened . . . [p. 539]" is to suppose that we know a priori, or by common sense, just what makes the system tick. But, as for releasers, we may guess and we may be wrong. Here it is a matter of measuring the depth given by various fin lengths (using luminous figures with no visible background) and correlating seen depth in this situation with the illusion; but regrettably this has not yet been done. Krueger's expectation is perhaps based on a misunderstanding of depth cue scaling. If we thought of the ends of the converging fins as "anchored" at a given

---

[1] Requests for reprints should be sent to Richard L. Gregory, Brain and Perception Laboratory, University Walk, University of Bristol, Bristol, England.

distance, then we indeed should expect increase in fin length to indicate increasing distance of the enclosed line—with resulting increase of illusory expansion—as the fins become longer, without limit. But this would be "hypothesis scaling," for it would occur as a function of apparent distance of the distorted enclosed line. The textured background prevents the figure's perspective features from serving to make it appear three-dimensional, so hypothesis scaling is not operating. This leaves only depth cue scaling in these conditions, but unfortunately we do not know independently of illusion data how depth cue scaling is related to the length of converging lines. In the world of normal objects, increase in length of receding lines is only associated with increase in distance when their near ends are fixed: since this is rarely the case, we should not expect this assumption to be built into depth cue scaling.

Krueger prefers assimilation, contrast, and figure/ground, as being "more peripheral and less cognitive." But on what theory of perception would this be accepted? A simple mechanism for cue scaling can readily be imagined; but surely assimilation, contrast (in this sense), and figure/ground seem to require far more elaborate "unconscious inference" than something like a releaser, set directly by certain stimulus features irrespective of the kind of object involved. I find these concepts too distantly related to the phenomena of illusory distortion to be relevant.

We are dealing here with perceptual scale distortions, so mechanisms capable, at least in principle, of producing size scale changes should be sought. Constancy scaling (which we know on very general grounds exists) is a candidate because it is clearly capable of modifying apparent size: either directly from depth cues or as a function of apparent distance.

Unlike depth cue scaling, depth hypothesis scaling *is* essentially cognitive, for it is tied to the apparent (or the prevailing hypothetical) distance of parts of figures or objects. We know that this is not set directly by features of the retinal image, because it varies as a function of apparent distance though the retinal image remains constant. The fact that hypothesis scaling is a clearly relevant cognitive phenomenon makes it especially interesting.

REFERENCES

GREGORY, R. L. Distortion of visual space as inappropriate constancy scaling. *Nature*, 1963, **119**, 678.

GREGORY, R. L. *The intelligent eye*. London: Weidenfeld & Nicolson, 1970.

KRUEGER, L. E. Gregory's theory of illusions: Some disconfirming evidence in the case of the Muller-Lyer illusion. *Psychological Review*, 1972, **79**, 538–539.

(Received April 19, 1972)

Psychological Review
1972, Vol. 79, No. 6, 542–546

ROLE OF THE HYPOTHALAMUS IN MOTIVATION: AN EXAMINATION OF VALENSTEIN'S REEXAMINATION

ERNEST H. BERGQUIST[1]

University of South Carolina

A recent (1970) article by E. S. Valenstein, V. C. Cox, and J. W. Kakolewski re-examining the role of the hypothalamus in motivation is reviewed and criticized. Alternative views regarding the anatomical specificity of hypothalamic mechanisms involved in the control of behavior are considered, and it is concluded that interdigitation theory is in closest agreement with available evidence. The question of whether stimulation may induce motivation (drive) is also discussed, and a generally affirmative decision reached. Implications of the fact that electrical stimulation is a gross and highly abnormal means of exciting brain mechanisms involved in the control of behavior are considered.

In a recent issue of this journal, Valenstein, Cox, and Kakolewski (1970) presented a reexamination of the role of the hypothalamus in motivation. The degree of anatomical specificity of stimulation-elicited behaviors and the nature of motivational conditions induced by such stimulation were considered. In the present article, attention is given to these same issues, and certain conclusions of Valenstein and his associates are criticized.

ANATOMICAL SPECIFICITY

Regarding the matter of anatomical specificity, there would seem to be the following four basic alternatives:

1. There may be distinct zones associated with the control of different traditional response categories (e.g., eating, drinking, sex, attack, threat).
2. There may be distinct zones associated with the control of broadly redefined response categories (where, using Valenstein's terminology, all responses in a given category may be associated with the same "general state," or "mood," of the organism).
3. There may be separate neural mechanisms mediating different response categories (either traditional or redefined), but with some anatomical interdigitation of these mechanisms.
4. The hypothalamus may be behaviorally equipotential.

The first alternative is definitely in error. As is brought out in the Valenstein article, there is considerable overlap among the anatomical sites that have been reported by various investigators to yield behaviors in different traditional response categories.

Valenstein and his associates clearly support either Alternative 2:

> Hypothalamic stimulation does not activate only one specific behavior pattern. The stimulation seems to excite the substrate for a group of responses that in a given species are related to a common state [Valenstein, Cox, & Kakolewski, 1970, p. 30].

or Alternative 4:

> The data have forced us to conclude that within the hypothalamus discrete anatomical regions cannot be found to correlate with the elicitation of any of the behaviors studied . . . [p. 20],

and

> anatomical location does not provide a good indicator of the behavior which will be elicited . . . [p. 25].

The present author thinks that Alternative 3 best handles the available evidence and attempts to show this in the current article.

RESPONSE SUBSTITUTION

Valenstein et al. were able to elicit commonly a second response after removing the initially preferred goal object without altering the original stimulation parameters. The responses they most frequently found to be interchangeable in this manner were eating, drinking, and gnawing, elicited mainly from the lateral hypothalamus. Either Alternative 2 or 3 could account for this observation, while Alternative 1 clearly cannot. Alternative 4

---

[1] Requests for reprints should be sent to Ernest H. Bergquist, Department of Psychology, University of South Carolina, Columbia, South Carolina.

(equipotentiality) would seem to predict that any and all behaviors should be elicitable from any hypothalamic site, something which Valenstein has not yet demonstrated.

SELECTIVE STIMULATION

The more selective stimulation techniques, such as thermal and chemical stimulation, provide a means of testing the validity of these various theories of hypothalamic function. Alternative 3 (interdigitation) would predict that the behavior elicited from a particular site might vary in kind depending on the stimulus agent, with different agents selectively activating different anatomically overlapping mechanisms. In contrast, Alternatives 2 and 4 would predict that the elicited behavior would be independent of the stimulating agent, and entirely determined by such factors as prior experience, availability of goal objects, and individual response prepotency. Available evidence appears to support the interdigitation theory.

It has been demonstrated in opossums, for example, that localized warming of the medial preoptic area using radio frequency current will evoke thermoregulatory responses such as grooming, panting, and sleeplike relaxation, but will not evoke nonthermoregulatory responses such as looking around, sniffing, locomotion, mouth opening, climbing, mating, attacking, and various stereotyped motor responses that can be elicited with 60-hertz sine wave current delivered through these same electrodes (Roberts, Bergquist, & Robinson, 1969). Localized warming of other hypothalamic regions was found to have no behavioral effect. This would seem to indicate that radio frequency current can selectively activate thermoregulatory response mechanisms localized in the medial preoptic area, while not activating other response mechanisms present in this same region, and clearly supports the interdigitation theory of hypothalamic function.

Also supporting the interdigitation view is Grossman's (1962) finding that eating can be evoked with an adrenergic substance, such as norepinephrine, and drinking with a cholinergic substance, such as carbachol, delivered via the same hypothalamic cannula. This finding strongly suggests that eating and drinking mechanisms may be closely intertwined with one another in the hypothalamus, but with the eating mechanism having adrenergic synaptic junctions and the drinking mechanism having cholinergic synaptic junctions.

Valenstein and his associates questioned this interpretation, suggesting that one, or both, of these chemical substances may be diffusing some distance away from the cannula tip before having its effect, thus leaving open the possibility that the two substances may be exciting anatomically distinct zones. This is unlikely. Various lines of evidence (behavioral, histochemical fluorescence, and radioactive tracing) all lead to the conclusion that these chemicals generally do not get more than one-half millimeter beyond the tip of the cannula (see Hoebel, 1971, pp. 552–553 for a summmary of this evidence).

If these effects are indeed both mediated by neural tissue near the tip of the cannula, an interesting test of the response substitution aspect of Valenstein's theory becomes possible. If adrenergic (cholinergic) substance is applied in the absence of food (water) at a site from which it has been verified that eating (drinking) can be evoked, Valenstein's theory would seem to predict that there may eventually be an increase in water (food) consumption, or perhaps in some cases an increased incidence of gnawing. While this particular experiment remains to be done, observations by Grossman (1962) suggest that aderenergic (cholinergic) stimulation, initially at least, may have an effect just opposite to this prediction, causing a decrease in water (food) consumption.

FURTHER ARGUMENTS AGAINST ALTERNATIVE 4

Alternative 4 (equipotentiality), although supported by the considerable overlap among anatomical sites reported to yield different behaviors, is contradicted by the specificity that has also been observed. In their mapping of the opossum hypothalamus, for example, Roberts, Steinberg, and Means (1967) found everything from nearly complete overlap of effective zones (e.g., mating and grooming) to no overlap at all (e.g., mating and threat).

Some specificity has recently been reported also regarding output pathways from hypothalamic mechanisms involved in the control of behavior. In an investigation of the effect of lesions at different loci on stimulation current thresholds for the elicitation of various behaviors in opossums, Bergquist (1970) found that lesions in the ventrolateral quadrant (especially those interrupting the medial forebrain bundle) of the ipsilateral posterior hypothalamus were most effective in elevating male mating, grooming, eating, attack, and investigatory thresholds, while lesions in the dorsomedial quadrant (particularly those interrupting posterior periventricular system fibers) were most effective in elevating thresholds for defensive threat. Similarly, Chi and Flynn (1971) reported that lesions made at sites from which quiet attack can be evoked in cats result in degeneration of medial forebrain bundle fibers, while lesions made at sites from which affective attack (similar in some respects to defensive threat in opossums) can be evoked result in degeneration of periventricular system fibers.

Also in opposition to the equipotentiality view is the behavioral specificity of hypothalamic lesions. Lateral lesions cause aphagia and adipsia (Teitelbaum & Epstein, 1962), ventromedial lesions cause hyperphagia and increased affective display (Hetherington & Ransom, 1940; Wheatley, 1944), anterior lesions in at least some species cause diminished sexual activity (Sawyer, 1960), and so on. Equipotentiality theory, in contrast, would predict that behavioral deficits would be unrelated to lesion locus.

FURTHER ARGUMENTS AGAINST ALTERNATIVE 2

Alternative 2 suggests that there may be hypothalamic zones involved in the control of broadly redefined response categories, with all responses in a given category being associated with a common organismic state; for example, eating, drinking, and gnawing (and possible other responses) might be controlled by a common substrate located in the lateral hypothalamus. The specific response elicited with stimulation of any such zone (out of the collection of potential responses mediated by that zone) would by entirely determined by such factors as prior experience, availability of goal objects, and individual response prepotency. It would be independent of both the type of stimulation used and the precise site of stimulation within the effective zone.

That stimulation-elicited behavior may be partially determined by factors such as prior experience, availability of goal objects, and individual response prepotency is not in question. What is being questioned in this article is the prediction that such behavior will be independent of the type of stimulation used and the specific site of stimulation. We have already shown, in fact (see section on selective stimulation), that elicited behavior can be quite dependent upon the kind of stimulation used. That specific site of stimulation may also be an important determining factor is indicated by a recent report (Olds, Allan, & Briese, 1971) that by using reduced surface-area-stimulating electrodes, eating and drinking behavior can be independently evoked from the lateral hypothalamus in the rat. This finding, while contrary to Alternative 2, is precisely what Alternative 3 would predict. Small surface-area-stimulating electrodes should be more likely to independently excite neural mechanisms that are immediately adjacent to, or interdigitated with, one another.

A further difficulty with Alternative 2 is that it may necessitate some rather improbable redefinition of response categories. For example, although mating, grooming, yawning, relaxing, pacing back and forth at the front of the test chamber, and thermoregulatory behavior have been evoked in the opossum from closely overlapping zones in the anterior hypothalamus and preoptic area (Roberts et al., 1967; Roberts et al., 1969), it is difficult to conceive of these behaviors as all being related to a common organismic state. It seems much more likely that these diverse behaviors are controlled by distinct mechanisms that simply happen to be in close physical proximity, or interdigitation, in this region.

STIMULATION-INDUCED MOTIVATION

The question of whether hypothalamic stimulation may induce motivation can with greatest clarity be treated in two parts. First, what is motivation (drive)? Second, does hypothalamic stimulation induce it?

The present author conceives of the term "motivation" as being a label referring to the activity of certain neural mechanisms, where any alteration in this activity (i.e., a change in motivation) will be reflected behaviorally in an altered likelihood of occurrence of certain behaviors and a change in reward value of certain related goal objects. Since hypothalamic stimulation has been found in many cases to have both of these effects (Anderson, Larson, & Persson, 1960; Roberts & Carey, 1965; Roberts & Kiess, 1964; Roberts et al., 1967), it can be concluded that such stimulation can induce motivation (drive), as so defined.

Unfortunately, Valenstein et al. did not clearly specify their definition of motivation. They did make clear, however, their view that motivational states are aversive. In referring to the question of whether hypothalamic stimulation might induce motivation, Valenstein et al. stated that this would require the postulation "that hypothalamic stimulation evokes specific aversive states associated with gnawing, food shuffling, tail preening, etc. [p. 20]". Since behavior eliciting hypothalamic stimulation is typically found to be rewarding, they concluded that such stimulation is apparently not inducing motivation.

This conclusion can be criticized in three different respects. First, it has not been established that motivation (drive) is necessarily always aversive. Some motivational conditions may well be aversive (e.g., 48-hour-food-deprivation-induced hunger), while others may not be (e.g., motivational conditions associated with sexual behavior, maternal behavior, tail preening, etc.).

Second, even if stimulation does induce an aversive motivational state, the total effect of that stimulation could still be positive if it simultaneously excites physically adjacent, or interdigitated, reward mechanisms. That such anatomical overlap of distinct motivational and reward mechanisms may actually occur in the hypothalamus is indicated by a recent report (Olds et al., 1971) that by using reduced surface-area-stimulating electrodes, motivational (eating or drinking) and reward effects can be independently evoked from areas (lateral hypothalamus) from which larger surface area electrodes commonly may evoke both effects.

Third, it is not true that behavior-eliciting stimulation is entirely nonaversive. Miller (1960), for example, reported that although his subjects would learn to press a lever to receive a brief pulse of stimulation to a site from which eating could be evoked, they would also learn a T maze to turn off maintained stimulation.

As further support for their view that hypothalamic stimulation does not induce motivation, Valenstein et al. pointed to their finding that stimulation may in some cases not enhance the reward value of all related goal objects equally (e.g., solid food pellets may be increased in reward value, but not liquid food), and in other cases may enhance the reward value of two or more unrelated goal objects (e.g., food pellets and water).

In interpreting these findings it is essential to bear in mind that electrical stimulation delivered via an implanted electrode is a highly abnormal means of activating hypothalamic control mechanisms. Because of this, mechanisms that are normally activated as a total unit may sometimes be only partially activated. As a result, incomplete response patterns may occur that might seldom or never occur naturally (see Roberts et al., 1967, for a detailed description of the variety of incomplete response patterns they observed with hypothalamic stimulation in the opossum), and, correspondingly, some goal objects may be enhanced in reward value, while other seemingly related goal objects may not be.

Similarly, two or more separate neural mechanisms may sometimes be simultaneously activated by stimulation (if, e.g., they are interdigitated with one another), even though this may never occur normally. As a result, combinations of two or more behavior patterns may occur that might seldom or never be observed naturally, and, correspondingly, two or more seemingly unrelated goal objects may be enhanced simultaneously in reward value.

In summary, the present author is suggesting that stimulation may oftentimes excite either fragments or combinations of specific motivational mechanisms. It is further suggested that this is why stimulation may in some cases not enhance the reward value of all related goal objects equally, and in other cases may enhance the reward value of two or more unrelated goal objects. This is in contrast to the view of Valenstein and his associates that stimulation is exciting mechanisms that mediate more "general states," or "moods" (see, e.g., Valenstein et al., p. 30). It is difficult to conceive, however, what kind of general state or mood would cause some of the combinations of reward effects reported (e.g., enhanced reward value of food pellets and water, but not liquid food).

REFERENCES

ANDERSSON, B., LARSON, S., & PERSSON, N. Some characteristics of the hypothalamic drinking center in the goat as shown by the use of permanent electrodes. *Acta Physiologica Scandinavica*, 1960, **50**, 140–152.

BERGQUIST, E. H. Output pathways of hypothalamic mechanisms for sexual, aggressive, and other motivated behaviors in opossum. *Journal of Comparative and Physiological Psychology*, 1970, **70**, 389–398.

CHI, C. C., & FLYNN, J. P. Neural pathways associated with hypothalamically elicited attack behavior in cats. *Science*, 1971, **171**, 703–706.

GROSSMAN, S. P. Direct adrenergic and cholinergic stimulation of hypothalamic mechanisms. *American Journal of Physiology*, 1962, **202**, 872–882.

HETHERINGTON, A. W., & RANSON, S. W. Hypothalamic lesions and adiposity in the rat. *Anatomical Record*, 1940, **78**, 149.

HOEBEL, B. G. Feeding: Neural control of intake. *Annual Review of Physiology*, 1971, **33**, 533–559.

MILLER, N. E. Motivational effects of brain stimulation and drugs. *Federation Proceedings*, 1960, **19**, 846–854.

OLDS, J., ALLAN, W. S., & BRIESE, E. Differentiation of hypothalamic drive and reward centers. *American Journal of Physiology*, 1971, **221**, 368–375.

ROBERTS, W. W., BERGQUIST, E. H., & ROBINSON, T. C. L. Thermoregulatory grooming and sleep-like relaxation induced by local warming of preoptic area and anterior hypothalamus in opossum. *Journal of Comparative and Physiological Psychology*, 1969, **67**, 182–188.

ROBERTS, W. W., & CAREY, R. J. Rewarding effect of performance of gnawing aroused by hypothalamic stimulation in the rat. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 317–324.

ROBERTS, W. W., & KIESS, H. O. Motivational properties of hypothalamic aggression in cats. *Journal of Comparative and Physiological Psychology*, 1964, **58**, 187–193.

ROBERTS, W. W., STEINBERG, M. L., & MEANS, L. W. Hypothalamic mechanism for sexual, aggressive, and other motivational behaviors in the opossum, Didelphis Virginiana. *Journal of Comparative and Physiological Psychology*, 1967, **64**, 1–15.

SAWYER, C. H. Reproductive behavior. In J. Field (Ed.), *Handbook of physiology*. (Section 1, Vol. 2) Baltimore: Williams & Wilkins, 1960.

TEITELBAUM, P., & EPSTEIN, A. N. The lateral hypothalamic syndrome: Recovery of feeding and drinking after lateral hypothalamic lesions. *Psychological Review*, 1962, **69**, 74–90.

VALENSTEIN, E. S., COX, V. C., & KAKOLEWSKI, J. W. Reexamination of the role of the hypothalamus in motivation. *Psychological Review*, 1970, **77**, 16–31.

WHEATLEY, M. D. The hypothalamus and affective behavior in cats. *Archives of Neurology and Psychiatry*, 1944, **52**, 296–316.

(Received January 27, 1972)

U. S. POSTAL SERVICE

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

(Act of August 12, 1970: Section 3685, Title 39, United States Code)

SEE INSTRUCTIONS ON PAGE 2 (REVERSE)

1. TITLE OF PUBLICATION: PSYCHOLOGICAL REVIEW

2. DATE OF FILING: 9/27/72

3. FREQUENCY OF ISSUE: BI-MONTHLY

4. LOCATION OF KNOWN OFFICE OF PUBLICATION *(Street, city, county, state, ZIP code) (Not printers)*: Prince & Lemon Sts., Lancaster, Lancaster, Pa. 17604

5. LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS *(Not printers)*: 1200 17th Street, N. W. Washington, D. C. 20036

6. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR

PUBLISHER *(Name and address)*: American Psychological Association, 1200 17th St., N.W., Washington, D. C. 20036

EDITOR *(Name and address)*: George Mandler, University of California, San Diego, La Jolla, California 92037

MANAGER EDITOR *(Name and address)*: Harold P. Van Cott, 1200 17th St., N. W., Washington, D. C. 20036

7. OWNER *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)*

NAME	ADDRESS
American Psychological Association	1200 17th St., N.W., Washington, D. C. 20036

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES *(If there are none, so state)*

NAME	ADDRESS
None	

9. FOR OPTIONAL COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES (Section 132.121, Postal Service Manual)

39 U. S. C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates."

In accordance with the provisions of this statute, I hereby request permission to mail the publication named in Item 1 at the reduced postage rates presently authorized by 39 U. S. C. 3626.

(Signature and title of editor, publisher, business manager, or owner)

10. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES *(Section 132.122, Postal Manual)*

(Check one)

The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes ☒ Have not changed during preceding 12 months ☐ Have changed during preceding 12 months *(If changed, publisher must submit explanation of change with this statement.)*

11. EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NUMBER OF COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A. TOTAL NO. COPIES PRINTED *(Net Press Run)*	12,615	12,892
B. PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	None	None
2. MAIL SUBSCRIPTIONS	11,243	11,418
C. TOTAL PAID CIRCULATION	11,243	11,418
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS 1. SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	107	110
2. COPIES DISTRIBUTED TO NEWS AGENTS, BUT NOT SOLD	None	None
E. TOTAL DISTRIBUTION *(Sum of C and D)*	11,350	11,528
F. OFFICE USE, LEFT-OVER, UNACCOUNTED, SPOILED AFTER PRINTING	1,265	1,364
G. TOTAL *(Sum of E & F—should equal net press run shown in A)*	12,615	12,892

I certify that the statements made by me above are correct and complete. ▶ Henry S. Salmon Jr Bus Mgr

(Signature of editor, publisher, business manager, or owner)

PS Form 3526 July 1971

Environment and the Social Sciences: perspectives and applications

represents a dawning realization that psychology has a substantial contribution to make to a better understanding of our present-day environmental problems. . .

The 23 chapters include investigations not only by psychologists, but by professionals in architecture, ecology, ethology, geography, resource management and planning, psychiatry, and sociology. Editors Joachim F. Wohlwill and Daniel H. Carson have grouped the chapters into three parts: general environmental conditions, special environmental settings, and environmental decision making.

This new book is based on a series of symposia, round table discussions, and paper sessions on environmental topics presented at the APA Annual Convention in September 1970. Additional material was solicited from specialists working on significant problems not dealt with at that convention, while a number of previously published papers were also selected to strengthen coverage of certain topics. Printed on recycled paper.

300 pages

All orders must be prepaid.

$5.50

ORDER FORM

_____ total enclosed for

_____ copies of *Environment and the Social Sciences: perspectives and applications.*

Name_____

Address _____

_____ Zip _____

Order Department x
AMERICAN PSYCHOLOGICAL ASSOCIATION
1200 Seventeenth Street, N.W.
Washington, D.C. 20036

Bind Your Journal Issues Into Valuable Vigilant Volumes

Pertinent information you sometimes desperately need is too often in the Journal issue you cannot find. Single copies have a way of getting lost, misplaced or destroyed. It is better for you to let us permanently bind each journal into semi-annual or annual volumes—then your reference source is always complete, organized and instantly at your service. We call them "Vigilant Volumes"; they so carefully store and provide on instant notice so much knowledge of timely value.

The cost per "Vigilant Volume," by official commission from the journal publishers is but $5.95 per volume, permanently hard bound in washable buckram, gold embossed with period dates, journal name and special insignia—plus your name stamped in gold leaf. "Vigilant Volumes" are handsome library additions too, real conversation pieces. These volumes are bound in the authorized colors.

HOW TO ORDER

Simply ship your journal issues to us via parcel post, together with your name, address, AND REMITTANCE, at $5.95 per volume. Within 45 days after receipt, bound volumes will be shipped to you. Price subject to change without notice. Please remit 50¢ per volume for return transportation and handling.

PUBLISHER'S AUTHORIZED
BINDERY SERVICE, Ltd.

(Authorized Binders of All Journals)
4440 W. Roosevelt Rd., Chicago, Illinois 60624

2 FEB 1974

An Esalen Book

The One Quest

by Claudio Naranjo

Complete human development. How does one approach it? Claudio Naranjo makes a broadly reasoned examination of the practices of religion, education, and psychotherapy over many lands and eras to illuminate the common ingredients of all forms in which people today search for ways of human growth.

Part I: An Introduction to the Quest for Growth
Part II: The Goals of Human Transformation
Part III: Traditional Ways and Contemporary Echoes
Part IV: The Oneness of Experience in the Ways of Growth
$8.95

Write for a free brochure describing Viking/Grossman publications in the field of psychology. **THE VIKING PRESS,** Dept. ATE-PR, 625 Madison Avenue, New York, N.Y. 10022

THE VIKING PRESS